

## METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS

### FIELD OF THE INVENTION

5

The present invention provides the three-dimensional structure of a histone acetyltransferase bromodomain. The three-dimensional structural information is included in the invention. The present invention also identifies for the first time, that bromodomains can bind to an acetylated binding partners. The interaction between bromodomains and their binding partners play a crucial role in various cellular functions, including in the regulation/modulation of DNA transcription. Therefore, the present invention provides procedures for identifying agents that can modulate the interaction of bromodomains and their binding partners by high throughput drug screening and/or through the use of rational drug design based on the three-dimensional data provided herein.

### BACKGROUND OF THE INVENTION

In recent years great strides have been made in the elucidation of the steps involved in intercellular and intracellular signaling. Indeed, the individual steps of the cascade of events involved in a number of cellular signal transduction processes have been determined. For example, intercellular signal transduction generally begins with an intercellular ligand binding the extracellular portion of a receptor of the plasma membrane. The bound receptor then either directly or indirectly initiates the activation of one or more cellular factors. An activated cellular factor may act as transcription factor by entering the nucleus to interact with its corresponding genomic response element, or alternatively, it may interact with other cellular factors depending on the complexity of the process. In either case, one or more transcription factors ultimately bind to one or more specific genomic response elements. This binding plays a crucial role in the up and/or down regulation of the transcription of the specific genes that are under the control of these genomic response elements. However, the process of re-organizing the chromatin of eukaryotic cells, which is a prerequisite for the binding of the transcription factor to the genomic response elements, has remained a mystery.

- Chromatin contains several highly conserved histone proteins including: H3, H4, H2A, H2B, and H1. These histone proteins package eukaryotic DNA into repeating nucleosomal units that are folded into higher-order chromatin fibers [Luger and
- 5 Richmond, *Curr. Opin. Genet. Dev.* **8**:140-146 (1998)]. A portion of the histone that comprises roughly a quarter of the protein protrudes from the chromatin surface, and is thereby sensitive to proteolytic enzymes [van Holde, in *Chromatin* (Rich, A., ed., Springer, New York ) pages 111-148 (1988); Hect *et al.*, *Cell* **80**:583-592 (1995)]. This portion of the histone is known as the “histone tail”. Histone tails tend to be free
- 10 for protein-protein interaction, and are also the portion of the histone most prone to post-translational modification. Such post-translational modification includes acetylation, phosphorylation, methylation, ubiquitination, and ADP-ribosylation [van Holde, in *Chromatin* (Rich, A., ed., Springer, New York ) pages 111-148 (1988)].
- 15 Of all classes of proteins, histones are amongst the most susceptible to post-translational modification. Perhaps the best studied post-translational modification of histones is the acetylation of specific lysine residues [Grunstein, M., *Nature*, **389**:349-352 (1997)]. Indeed, acetylation of histone lysine residues has been suggested to play a pivotal role in chromatin remodeling and gene activation. Consistently,
- 20 distinct classes of enzymes, namely histone acetyltransferases (HATs) and histone deacetylases (HDACs), acetylate or de-acetylate specific histone lysine residues [Struhl, *Genes Dev.* **12**:599-606 (1998)].

- Nearly all known nuclear HATs contain an approximately 110 amino acid sequence
- 25 known as the bromodomain [Jeanmougin *et al.*, *Trends in Biochemical Sciences*, **22**:151-153 (1997)], a protein motif that was initially discovered in *Drosophila* brahma protein. Bromodomains are found in a large number of chromatin-associated proteins and have now been identified in approximately 40 proteins, often adjacent to other protein motifs [Jeanmougin *et al.*, *Trends in Biochemical Sciences*, **22**:151-153 (1997); Tamkun *et al.*, *Cell*, **68**:561-572 (1992); Hanes *et al.*, *Nucleic Acids Research*, **20**:2603 (1992)]. Proteins that contain a bromodomain often contain a second
- 30 bromodomain. However, despite the wide occurrence of bromodomains and their

likely role in chromatin regulation, their three-dimensional structure and binding partners heretofore have remained unknown.

Therefore, there is a need to identify a binding partner for a bromodomain. In addition, there is a need to identify agonists or antagonists to the bromodomain-binding partner complex. Since a preferred method of drug-screening relies on structure based drug design, there is also a need to determine the three-dimensional structure of a bromodomain. In this case, once the three dimensional structure of bromodomain is determined, potential agonists and/or potential antagonists can be designed with the aid of computer modeling [Bugg *et al.*, *Scientific American*, Dec.:92-98 (1993); West *et al.*, *TIPS*, 16:67-74 (1995); Dunbrack *et al.*, *Folding & Design*, 2:27-42 (1997)]. However, heretofore the three-dimensional structure of the bromodomain has remained unknown. Therefore, there is a need for obtaining a form of the bromodomain that is amenable for NMR analysis and/or X-ray crystallographic analysis. Furthermore, there is a need for the determination of the three-dimensional structure of the bromodomain. Finally, there is a need for procedures for related structural based drug design predicated on such structural data.

The citation of any reference herein should not be construed as an admission that such reference is available as "Prior Art" to the instant application.

### SUMMARY OF THE INVENTION

The present invention provides, for the first time, that bromodomains bind to acetyl-lysine residues of proteins. The present invention also provides the three-dimensional structure of a bromodomain as well as the three-dimensional structure of a bromodomain-acetyl-histamine complex. The structural information provided can be employed in methods of identifying drugs that can modulate the cellular processes that involve bromodomain-acetyl-lysine interactions. These interactions include chromatin remodeling, which is a required step in eukaryotic transcription. In a particular embodiment, the three-dimensional structural information is used in the design of an inhibitor of leukemia.

The present invention provides an isolated nucleic acid that encodes a peptide consisting of about 21 to 40 amino acids that comprises a ZA loop of a bromodomain. In a preferred embodiment the peptide comprises about 23 to 34 amino acids. The isolated nucleic acid can further comprise a heterologous nucleotide sequence.

5

In a preferred embodiment the peptide comprises the amino acid sequence of SEQ ID NO:3. In another embodiment the peptide comprises the amino acid sequence of SEQ ID NO:43. In particular embodiments the ZA loop is obtained from the bromodomain having the amino acid sequence of SEQ ID NO:7, or SEQ ID NO:8, or SEQ ID NO:9, or SEQ ID NO:10, or SEQ ID NO:11, or SEQ ID NO:12, or SEQ ID NO:13, or SEQ ID NO:14, or SEQ ID NO:15, or SEQ ID NO:16, or SEQ ID NO:17, or SEQ ID NO:18, or SEQ ID NO:19, or SEQ ID NO:20, or SEQ ID NO:21, or SEQ ID NO: 22, or SEQ ID NO:23, or SEQ ID NO:24, or SEQ ID NO:25, or SEQ ID NO:26, or SEQ ID NO:27, or SEQ ID NO:28, or SEQ ID NO:29, or SEQ ID NO:30, or SEQ ID NO: 15 or SEQ ID NO:31, or SEQ ID NO:32, or SEQ ID NO: 33, or SEQ ID NO:34, or SEQ ID NO:35, or SEQ ID NO:36 , or SEQ ID NO:37, or SEQ ID NO:38, or SEQ ID NO: or SEQ ID NO:39, or SEQ ID NO:40, or SEQ ID NO:41, or SEQ ID NO:42.

The present invention further provides a recombinant DNA molecule that comprises an isolated nucleic acid of the present invention, as described above, with or without a heterologous nucleotide sequence. Such a recombinant DNA molecule can be operatively linked to an expression control sequence and can be part of an expression vector. The present invention further provides a cell that comprises such an expression vector. The cell can be either a eukaryotic or a prokaryotic cell. The present invention further provides a method of expressing the peptides of the present invention or fragments thereof in this cell. One such method comprises culturing the cell in an appropriate cell culture medium under conditions that provide for expression of the peptide by the cell.

30 The present invention further provides a peptide consisting of about 21 to 40 amino acids that comprises a ZA loop of a bromodomain. In a preferred embodiment the



peptide comprises about 23 to 34 amino acids. The present invention also provides fusion proteins or peptides comprising these peptides.

In a preferred embodiment the peptide comprises the amino acid sequence of SEQ ID NO:3. In another embodiment the peptide comprises the amino acid sequence of SEQ ID NO:43. In particular embodiments the ZA loop is obtained from the bromodomain having the amino acid sequence of SEQ ID NO:7, or SEQ ID NO:8, or SEQ ID NO:9, or SEQ ID NO:10, or SEQ ID NO:11, or SEQ ID NO:12, or SEQ ID NO:13, or SEQ ID NO:14, or SEQ ID NO:15, or SEQ ID NO:16, or SEQ ID NO:17, or SEQ ID NO:18, or SEQ ID NO:19, or SEQ ID NO:20, or SEQ ID NO:21, or SEQ ID NO: 22, or SEQ ID NO:23, or SEQ ID NO:24, or SEQ ID NO:25, or SEQ ID NO:26, or SEQ ID NO:27, or SEQ ID NO:28, or SEQ ID NO:29, or SEQ ID NO:30, or SEQ ID NO: or SEQ ID NO:31, or SEQ ID NO:32, or SEQ ID NO: 33, or SEQ ID NO:34, or SEQ ID NO:35, or SEQ ID NO:36 , or SEQ ID NO:37, or SEQ ID NO:38, or SEQ ID NO: or SEQ ID NO:39, or SEQ ID NO:40, or SEQ ID NO:41, or SEQ ID NO:42.

The present invention also provides antibodies raised against the peptides/proteins of the present invention, or raised against an antigenic fragment of these proteins/fragments. In a particular embodiment an antibody is raised against a fragment of the ZA loop of a bromodomain. In another embodiment an antibody is raised against a fragment of a protein or peptide that comprises an acetyl-lysine, wherein the protein or peptide can bind to a bromodomain. Such fragments can be conjugated to a carrier protein or be part of a fusion protein. In one embodiment the antibody is a polyclonal antibody. In another embodiment, the antibody is a monoclonal antibody. A hybridoma that makes the monoclonal antibody is also part of the present invention. In a particular embodiment the antibody is a chimeric antibody. Antibodies that can specifically recognize acetyl-lysine residues involved bromodomain binding are also part of the present invention.

In another aspect of the present invention a method is provided for identifying a compound that modulates the affinity of a bromodomain for a ligand (and/or protein) that comprises an acetylated lysine. One such embodiment comprises contacting the

bromodomain and the ligand in the presence of a compound under conditions that ,  
the bromodomain and the ligand bind in the absence of the compound. The affinity of  
the bromodomain for the ligand is then determined (*e.g.*, measured). A compound is  
identified as a compound that modulates the affinity of the bromodomain for the  
5 ligand when there is a change in the affinity of the bromodomain for the ligand in the  
presence of the compound. When the affinity of the bromodomain for the ligand  
increases in the presence of the compound, the compound is identified as a promoting  
agent for the bromodomain-ligand complex. When the affinity of the bromodomain  
for the ligand decreases in the presence of the compound, the compound is identified  
10 as an inhibitor of the bromodomain-ligand complex. In a preferred embodiment, the  
compound to be tested is pre-selected by performing rational drug design with the set  
of atomic coordinates obtained from one or more of Tables 1-6. More preferably the  
selecting is performed in conjunction with computer modeling. In a particular  
embodiment, the compound is selected by performing rational drug design with the  
15 set of atomic coordinates obtained from a set of atomic coordinates defining the three-  
dimensional structure of a bromodomain consisting of the amino acid sequence of  
SEQ ID NO:7 alone or with acetyl-histamine.

The present invention also provides a method of identifying a compound that  
20 modulates the stability of a bromodomain-acetyl-lysine binding complex. One such  
embodiment comprises contacting the bromodomain-acetyl-lysine binding complex in  
the presence of the compound under conditions in which the bromodomain-acetyl-  
lysine binding complex forms in the absence of the compound. The stability of the  
bromodomain-acetyl-lysine binding complex is then determined (*e.g.*, measured). A  
25 compound is identified as a compound that modulates the stability of the  
bromodomain-acetyl-lysine binding complex, when there is a change in the stability  
of the bromodomain-acetyl-lysine binding complex in the presence of that compound.  
When the stability of the bromodomain-acetyl-lysine binding complex increases in the  
presence of the compound, the compound is identified as a stabilizing agent. When  
30 the stability of the bromodomain-acetyl-lysine binding complex decreases in the  
presence of the compound, the compound is identified as an inhibitor. In a preferred  
embodiment, the compound to be tested is pre-selected by performing rational drug

design with the set of atomic coordinates obtained from one or more of Tables 1-6. More preferably the selecting is performed in conjunction with computer modeling. In a particular embodiment, the compound is selected by performing rational drug design with the set of atomic coordinates obtained from a set of atomic coordinates  
5 defining the three-dimensional structure of a bromodomain consisting of the amino acid sequence of SEQ ID NO:7 alone or with acetyl-histamine.

As anyone having skill in the art of drug development would readily understand, the potential drugs selected by the above methodologies can be refined by re-testing in  
10 appropriate drug assays, including those disclosed herein. Chemical analogs of such potential drugs can be obtained (either through chemical synthesis or drug libraries) and be analogously tested. Therefore, methods comprising successive iterations of the steps of the individual drug assays, as exemplified herein, using either repetitive or different binding studies, or transcription activation studies or other such studies are  
15 envisioned in the present invention. In addition, potential drugs may be identified first by rapid throughput drug screening, as described below, prior to performing computer modeling on a potential drug using the three-dimensional structure of the bromodomain.

20 The present invention further comprises all of the potential, selected, and putative compounds (drugs) identified by the methods of the present invention, as well as the final drugs themselves identified with the methods of the present invention.

The present invention further provides a method for identifying a potential binding  
25 partner for a protein (*e.g.*, a histone) comprising an acetyl-lysine. One such embodiment comprises contacting the protein with a polypeptide comprising a bromodomain. In a preferred embodiment the bromodomain comprises the amino acid sequence of SEQ ID NO:3. In particular embodiments the bromodomain has the amino acid sequence of SEQ ID NO:7, or SEQ ID NO:8, or SEQ ID NO:9, or SEQ ID  
30 NO:10, or SEQ ID NO:11, or SEQ ID NO:12, or SEQ ID NO:13, or SEQ ID NO:14, or SEQ ID NO:15, or SEQ ID NO:16, or SEQ ID NO:17, or SEQ ID NO:18, or SEQ ID NO:19, or SEQ ID NO:20, or SEQ ID NO:21, or SEQ ID NO: 22, or SEQ ID

NO:23, or SEQ ID NO:24, or SEQ ID NO:25, or SEQ ID NO:26, or SEQ ID NO:27,  
 or SEQ ID NO:28, or SEQ ID NO:29, or SEQ ID NO:30, or SEQ ID NO: or SEQ ID  
 NO:31, or SEQ ID NO:32, or SEQ ID NO: 33, or SEQ ID NO:34, or SEQ ID NO:35,  
 or SEQ ID NO:36 , or SEQ ID NO:37, or SEQ ID NO:38, or SEQ ID NO: or SEQ ID  
 5 NO:39, or SEQ ID NO:40, or SEQ ID NO:41, or SEQ ID NO:42.

The present invention further provides a method for identifying a protein having a  
 bromodomain. One such embodiment comprises contacting a cellular extract with a  
 peptide comprising an acetyl-lysine.

10

The present invention further provides agents that can inhibit the binding of a  
 bromodomain with a protein comprising an acetyl-lysine. In one embodiment the  
 agent is ISYGR-*AcK*-KRRQRR (SEQ ID NO:4). In another embodiment the agent is  
 ARKSTGG-*AcK*-APRKQL (SEQ ID NO:5). In still another embodiment the agent  
 15 is QSTSRHK-*AcK*-LMFKTE (SEQ ID NO:6). In yet another embodiment the agent  
 is an analog of acetyl-lysine such as acetyl-histamine. In still another embodiment the  
 agent is an antibody that recognizes an acetyl-lysine of a protein binding partner of a  
 bromodomain. In a preferred embodiment the agent is an antibody raised against a  
 ZA loop of a bromodomain. These agents can be used as pharmaceuticals in  
 20 compositions that contain a pharmaceutically acceptable carrier for example, or in the  
 various drug assays of the present invention, serving as controls to demonstrate  
 specificity.

Accordingly, it is a principal object of the present invention to provide the three-  
 25 dimensional coordinates of a bromodomain.

It is a further object of the present invention to provide the three-dimensional  
 coordinates of a bromodomain complexed with acetyl-histamine.

30 It is a further object of the present invention to provide an assay for identifying  
 proteins that contain bromodomains that bind proteins that comprise acetyl-lysine.

It is a further object of the present invention to provide methods of identifying drugs that can modulate the bromodomain-acetyl-lysine binding complex.

It is a further object of the present invention to provide methods of identifying drugs  
5 that can inhibit the binding of a bromodomain to a protein containing acetyl-lysine.

It is a further object of the present invention to provide methods that incorporate the use of rational design for identifying such drugs.

10 It is a further object of the present invention to provide a method of identifying drugs that can treat leukemia.

It is a further object of the present invention to provide a method of identifying drugs that can treat and/or prevent AIDS.

15

These and other aspects of the present invention will be better appreciated by reference to the following drawings and Detailed Description.

#### BRIEF DESCRIPTION OF THE DRAWINGS

20

Figure 1. Structure-based sequence alignment of a selected number of bromodomains. The sequences were aligned based on the NMR-derived structure of the P/CAF bromodomain, and the predicated four  $\alpha$ -helices are shown in green boxes.

Bromodomains are grouped on the basis of the sequence and/or functional similarities  
25 as described by Jeanmougin *et al.*, [Trends in Biochemical Sciences, 22:151-153 (1997)]. Residue numbers of the P/CAF bromodomain are indicated above its sequence. Three absolutely conserved residues, corresponding to Pro751, Pro767, and Asn803 in the P/CAF bromodomain, are shown in red. Highly conserved residues are colored in blue. The residues of the P/CAF bromodomain that interact with  
30 acetyl-histamine, as determined by intermolecular NOEs, are indicated by asterisks. The ZA loop, which is critical for acetyl-lysine binding, for each of the indicated bromodomains is also identified. The underlined residues were changed individually

by site-directed mutagenesis to Ala. Genbank accession numbers for the proteins are as indicated in Table 8, in the Example below, along with the SEQ ID NOs. for the bromodomain sequences.

- 5 Figures 2A-2H depict the structure of the P/CAF bromodomain. Figures 2A-2B shows the stereoview of the C $\alpha$  trace of 30 superimposed NMR-derived structures of the bromodomain (residues 722-830). The N-terminal four residues (SKEP) which are structurally disordered are omitted for clarity. For the final 30 structures, the root-mean-square deviations (RMSDs) of the backbone and all heavy atoms are 0.63  
10  $\pm 0.11$  Å and  $1.15 \pm 0.12$  Å for residues 723-830, respectively. The RMSDs of the backbone and all heavy atoms for the four  $\alpha$ -helices (residues 727-743, 770-776, 785-802, and 807-827), are  $0.34 \pm 0.04$  Å and  $0.87 \pm 0.06$  Å, respectively. Figures 2C-2D show the stereoview of the bromodomain structures from the bottom of the protein, which is rotated approximately 90° from the orientation in Figures 2A-2B.  
15 Figure 2E shows the Ribbons [Carson, M., *J. Appl. Crystallogr.* 24:958-961 (1991)] depiction of the averaged minimized NMR structure of the P/CAF bromodomain. The orientation of Figure 2E is as shown in Figures 2A-2B. Figures 2F-2G are schematic representations of the overall topology of the up-and-down four-helix bundle folds with the opposite handedness. The left-handed fold is seen in  
20 bromodomain, cytochrome *b<sub>5</sub>*, and T4 lysozyme (left, Figure 2F), whereas the right-handed four-helix bundles are observed in proteins such as hemerythrin and cytochrome *b<sub>562</sub>* (right, Figure 2G) [Richardson, J., *Adv. Protein Chem.*, 34:167-339 (1989); Presnell and Cohen, *Proc. Natl. Acad. Sci. USA* 86:6592-6596 (1989)]. Figure 2H is a molecular surface representation of the electrostatic potential (blue =  
25 positive; red = negative) of the bromodomain calculated in GRASP [Nicholls *et al.*, *Biophys. J.* 64:166-170 (1993)]. The hydrophobic and aromatic residues (Tyr809, Tyr802, Tyr760, Ala757, and Val752) located between the ZA and BC loops are indicated.
- 30 Figures 3A-3C show the binding of the P/CAF bromodomain to AcK. Figure 3A shows the superimposed region of the 2D <sup>15</sup>N-HSQC spectra of the bromodomain (approximately 0.5 mM) in its free form (red) and complexed to the AcK-containing

H4 peptide (molar ratio 1:6) (black). Figure 3B is the Ribbon and dotted-surface diagram of the bromodomain depicting the location of the lysine-acetylated H4 peptide binding site. The color coding reflects the chemical shift changes ( $\Delta\delta$ ) of the backbone amide  $^1\text{H}$  and  $^{15}\text{N}$  resonances upon binding to the AcK peptide as observed in the  $^{15}\text{N}$ -HSQC spectra. The normalized weighted average of the chemical shift changes was calculated by  $\Delta_{av}/\Delta_{max} = [\Delta\delta_{\text{NH}}^2 + \Delta\delta_{\text{N}}^2/25]/2^{1/2}/\Delta_{max}$ , where  $\Delta_{max}$  is the maximum weighted chemical shift difference observed for Tyr809 (0.16ppm). The backbone atoms are color-coded in red, yellow, or green for residues that have  $\Delta_{av}/\Delta_{max}$  of >0.6 (Tyr809, Glu808, Asn803, and Ala757), 0.2-0.6 (Ala813, Tyr802, Tyr760, and Val752), or <0.2 (Cys812, Ser807, Cys799, Phe796, and Phe748), respectively. The non-perturbed residues are shown in blue. Figure 3C shows the chemical structures of acetyl-lysine, acetyl-histamine, and acetyl-histidine.

Figure 4 depicts the acetyl-lysine binding pocket. This is the Ribbons [Carson, M., *J. Appl. Crystallogr.* **24**:958-961 (1991)] depiction of a portion of the P/CAF bromodomain complexed with the acetyl-histamine. The ligand is color-coded by atom type.

#### DETAILED DESCRIPTION OF THE INVENTION

20

The present invention identifies a general binding partner (ligand) for the protein motif known as the bromodomain. Indeed, by combining structural and site-directed mutagenesis studies the present invention demonstrates that bromodomains can interact specifically with acetyl-lysine (AcK), making them the first protein modules known to exhibit such interactions. Like other modular domains, such as Src homology-2 (SH2) and phosphotyrosine binding (PTB) domains, which specifically interact with phosphotyrosine-containing proteins, the bromodomain/acetyl-lysine recognition provides a means to regulate protein-protein interactions via protein lysine acetylation. The nature of the acetyl-lysine recognition by the bromodomain is similar to that of histone acetyltransferase interaction with acetyl-CoA. The present invention therefore couples for the first time, the functionality of the bromodomain with the HAT activity of coactivators in the regulation of gene transcription.

30

The present invention further provides both a nuclear magnetic resonance (NMR) structure of the bromodomain from the HAT coactivator P/CAF (p300/CBP-associated factor) as well as the structure for the P/CAF bromodomain in complex with acetyl-histamine. The structure reveals an unusual left-handed  
 5 up-and-down four-helix bundle.

The results disclosed herein explain prior deletion experiments which showed that the bromodomain is indispensable for the function of GCN5 in yeast.

Bromodomain-AcK binding also appears to be important for the assembly and activity  
 10 of multiprotein complexes in transcriptional activation. The results reported herein therefore, form the foundation for identifying specific biological ligands and for defining the molecular mechanisms by which the extensive family of bromodomains participate in chromatin remodeling and transcriptional activation

15 As disclosed herein, the binding partner for the bromodomain is a peptide or protein comprising an acetyl-lysine (AcK). Interestingly, whereas a free acetyl-lysine does not appear to bind the bromodomain, an analog of the acetyl-lysine, acetyl-histamine, does. This is most likely due to the additional charge present in the free amino acid. Consistently, free acetyl-histidine also does not to bind the bromodomain.

20

The present invention further provides a key region of the bromodomain for the interaction with its acetyl-lysine binding partner, the ZA loop. The amino acid sequence of the ZA loop is defined in Figure 1 for a number of bromodomains and is depicted in Figure 2A for P/CAF. In a particular embodiment, the ZA loop has  
 25 between about 21 and 40 amino acid residues comprising the amino acid sequence of :

F X<sub>2-3</sub> P X<sub>5-8</sub> J<sub>P/K/H</sub> X Y J<sub>Y/F/H</sub> X<sub>5</sub> P J<sub>M/I/V</sub> D (SEQ ID NO:3)

more preferably the ZA loop has about 23 to 34 amino acid residues and comprises the  
 30 amino acid sequence:

X<sub>2</sub> F X<sub>2-3</sub> P X<sub>5-8</sub> J<sub>P/K/H</sub> X Y J<sub>Y/F/H</sub> X<sub>5</sub> P J<sub>M/I/V</sub> D (SEQ ID NO:43)



(1) The single letter amino acid code is used in this description, *i.e.*, “F” for phenylalanine; “P” for proline; “Y” for tyrosine; and “D” for aspartic acid.

(2) “X” indicates any amino acid (an undesigned amino acid); and X, X<sub>2</sub>, X<sub>2-3</sub>, X<sub>5</sub>, and X<sub>5-8</sub> indicates one undesigned amino acid, two consecutive undesigned amino acids, two or three consecutive undesigned amino acids, five consecutive  
 5 undesigned amino acids, and five to eight consecutive undesigned amino acids respectively.

(3) “J” indicates that identity of the amino acid is restricted to a particular group, again the one letter code is used

- 10 :
- (i) J<sub>P/K/H</sub> is either proline, lysine or histidine.
  - (ii) J<sub>Y/F/H</sub> is either tyrosine, phenylalanine or histidine.
  - (iii) J<sub>M/I/V</sub> is either methionine, isoleucine, or valine.

Since this region of the bromodomain is important in binding its acetyl-lysine binding  
 15 partner, antibodies specifically raised against this region are also included in the present invention. In a particular embodiment, the antibody is a humanized chimeric antibody that can be used in therapeutic treatment. Thus monoclonal, chimeric, and polyclonal antibodies raised against bromodomains, preferably against amino acid residues in the ZA loop region are part of the present invention. In a specific  
 20 embodiment the antibody is raised against a peptide, fusion peptide or conjugated peptide consisting of amino acid residues 746 to 765 of SEQ ID NO:2, *i.e.*, WPFMEPVKRTEAPGYEYEVIR (SEQ ID NO:44). Such antibodies can be used in the treatment of leukemia for example. Alternatively, these antibodies can be used in drug discovery assays.

25

Thus the present invention provides the first detailed structural information regarding a bromodomain and a bromodomain complexed with its acetylated binding partner. The present invention therefore provides the three-dimensional structure of the bromodomain and a bromodomain acetylated binding partner complex. Since the  
 30 interaction of the bromodomain with a histone for example, can play a significant role in chromatin remodeling/regulation, the structural information provided herein can be employed in methods of identifying drugs that can modulate basic cell processes by modulating the transcription. In a particular embodiment, the three-dimensional

structural information is used in the design of a small organic molecule for the treatment of cancer.

Indeed, the bromodomain and lysine-acetylated protein interaction can now be  
5 implicated to play a causal role in the development of a number of diseases including  
cancers such as leukemia. For example, chromatin remodeling plays a central role in  
the etiology of viral infection and cancer [Archer and Hodin, *Curr. Opin. Genet. Biol.*  
9:171-174 (1999); Jacobson and Pillus, *Curr. Opin. Genet. Biol.* 9:175-184 (1999)].  
Both altered histone acetylation/deacetylation and aberrant forms of chromatin-  
10 remodeling complexes are associated with human diseases. Furthermore,  
chromosomal translocation of various cellular genes with those encoding HATs and  
subunits of chromatin remodeling complexes have been implicated in leukomogenesis.  
The *MOZ* (monocytic leukemia zinc finger) and *MLL/ALL-1* genes are frequently fused  
to the gene encoding the co-activator HAT CBP [Sobulo *et al.*, *Proc. Natl. Acad. Sci.*  
15 *USA* 94:8732-8737(1997)]. The resulting fusion protein MLL-CBP contains the  
tandem bromodomain-PHD finger-HAT domain of CBP. It also has been shown that  
both the bromodomain and HAT domain of CBP are required for leukomogenesis,  
because deletion of either the bromodomain or the HAT domain results in loss of the  
MLL-CBP fusion protein's ability for cell transform. These results indicate that the  
20 CBP bromodomain, and more particularly, the ZA loop of the CBP bromodomain, is  
an excellent target for developing drugs that interfere with the bromodomain acetyl-  
lysine interaction that can be used in the treatment of human acute leukemia. In  
addition, an antibody (*e.g.*, a humanized antibody) raised specifically against a peptide  
from the ZA loop of the CBP bromodomain could also be effective for treating these  
25 conditions.

Furthermore, the human immunodeficiency virus type 1 (HIV-1) *trans*-activator  
protein, Tat, is absolutely required for productive HIV viral replication [Jeang and  
Gatignol, *Curr. Top. Microbiol. Immunol.*, 188:123-144(1994)]. Recently, it has been  
30 shown that HIV-1 Tat transcriptional activity is tightly regulated by lysine acetylation  
[Kiernan *et al.*, *EMBO Journal* 18:6106-6118 (1999)]. Therefore, the interaction of  
the acetyl-lysine of Tat with one or more bromodomain-containing proteins associated

with chromatin remodeling could mediate gene transcription. Thus, the bromodomain/lysine-acetylated Tat interaction could also serve as a drug target for blocking HIV replication in cells. Similarly, an antibody raised specifically against a peptide from the ZA loop of the bromodomain could also be effective for treating these  
5 conditions.

In addition, based on the new structural information disclosed herein, the key amino acid residues for the binding of a given bromodomain and its binding partner can be identified and further elucidated using basic mutagenesis and standard isothermal  
10 titration calorimetry, for example. In this case, both the crucial amino acids for the bromodomain and the binding partner (i.e., apart from the acetyl-lysine) can be readily determined and are also part of the present invention.

The results obtained from the structural and functional studies disclosed herein provide  
15 the foundation for both high throughput drug screening and structure-based rational drug design. The agents identified by this procedure will be useful for ameliorating conditions involving chromatin remodeling/regulation as indicated above.

Structure based rational drug design is the most efficient method of drug development.  
20 However, heretofore, no information has been disclosed regarding the structure of the bromodomain or more importantly, its interaction with the acetyl-lysine of its binding partner. Obtaining detailed structural information requires an extensive NMR or X-ray crystallographic analysis. By determining and then exploiting the detailed structural information of the bromodomain and of the bromodomain/acetyl-histamine  
25 (exemplified by NMR analysis below) the present invention provides novel methods for developing new drugs through structure based rational drug design.

Thus the present invention provides representative sets of the atomic structure coordinates of the free form of the P/CAF bromodomain (Table 5) and of the P/CAF  
30 bromodomain-acetyl-histamine complex (Table 6) which were both obtained by NMR analysis. A Ribbon diagram of the three-dimensional structure of the P/CAF bromodomain is depicted in Figure 2E, whereas the P/CAF bromodomain acetyl-lysine

binding pocket is depicted in Figure 4. The present invention also provides the NOE-derived distance restraints, and NMR chemical shift assignments of the P/CAF bromodomain. The NMR chemical shift assignments of the P/CAF bromodomain are included in the chemical shift table (Table 1) for the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of P/CAF bromodomain. The unambiguous NOE-derived Inter-proton Distance Restraints (Table 2), the ambiguous NOE-derived Inter-proton Distance Restraints (Table 3) and the  $^1\text{H}$  bonding restraints (Table 4) are also disclosed herein. The sample atomic coordinate data provided enable the skilled artisan to practice the invention. In addition, Tables 1-6 are also capable of being placed into a computer readable form which is also part of the present invention. Furthermore, methods of using these coordinates and chemical shifts and related information (including in computer readable forms) either individually or together in drug assays are also provided. More particularly, such atomic coordinates can be used to identify potential ligands or drugs which will modulate the binding of a bromodomain with its binding partner.

Therefore, if appearing herein, the following terms shall have the definitions set out below.

As used herein a "bromodomain-acetyl-lysine binding complex" is a binding complex between a bromodomain or fragment thereof and either a peptide/polypeptide comprising an acetyl-lysine (or an analog of acetyl-lysine), or a free analog of acetyl-lysine, such as acetyl-histamine disclosed in the Example below. Preferably, the peptide comprises at least six amino acids in addition to the acetyl-lysine. The dissociation constant of a bromodomain-acetyl-lysine binding complex is dependent on whether the lysine residue or analog thereof is acetylated or not, such that the affinity for the bromodomain and the peptide comprising the lysine residue (for example) significantly decreases when that lysine residue is not acetylated.

As used herein a "ZA loop" of a bromodomain is one portion of a bromodomain that is involved in the binding of the bromodomain to the acetyl-lysine. The structure of the ZA loop of the bromodomain of for P/CAF is depicted in Figure 2A. The ZA loop has between about 20 and 40 amino acids and comprises the amino acid sequence of SEQ ID NO:3. More preferably the ZA loop comprises between about 23 to 34 amino acids

and has the amino acid sequence SEQ ID NO:43. The amino acid sequence of the ZA loop for a representative number of individual bromodomains is shown in Figure 1.

5 A "polypeptide" or "peptide" comprising a fragment of a bromodomain, such as the ZA loop, or a peptide or polypeptide comprising an acetyl-lysine, as used herein can be the "fragment" alone, or a larger chimeric or fusion peptide/protein which contains the "fragment".

As used herein the terms "fusion protein" and "fusion peptide" are used  
10 interchangeably and encompass "chimeric proteins and/or chimeric peptides" and fusion "intein proteins/peptides". A fusion protein comprises at least a portion of a protein or peptide of the present invention, *e.g.*, a bromodomain, joined *via* a peptide bond to at least a portion of another protein or peptide including *e.g.*, a second bromodomain in a chimeric fusion protein. In a particular embodiment the portion of  
15 the bromodomain is antigenic. Fusion proteins can comprise a marker protein or peptide, or a protein or peptide that aids in the isolation and/or purification of the protein, for example.

As used herein, and unless otherwise specified, the terms "agent", "potential drug",  
20 "compound", "test compound" or "potential compound" are used interchangeably, and refer to chemicals which potentially have a use as an inhibitor or activator/stabilizer of bromodomain-acetyl-lysine binding. Therefore, such "agents", "potential drugs", "compounds" and "potential compounds" may be used, as described herein, in drug assays and drug screens and the like.

25 As used herein a "small organic molecule" is an organic compound, including a peptide [or organic compound complexed with an inorganic compound (*e.g.*, metal)] that has a molecular weight of less than 3 Kilodaltons. Such small organic molecules can be included as agents, etc. as defined above.

30 As used herein the term "binds to" is meant to include all such specific interactions that result in two or more molecules showing a preference for one another relative to some third molecule. This includes processes such as covalent, ionic, hydrophobic and

hydrogen bonding but does not include non-specific associations such as solvent preferences.

As used herein the term "about" signifies that a value is within twenty percent of the indicated value *i.e.*, a peptide containing "about" 20 amino acid residues can contain between 16 and 24 amino acid residues.

General Techniques for Constructing Nucleic Acids That Encode the Bromodomains and Fragments Thereof (Including, ZA Loops); and the Bromodomain Binding Partners of the Present Invention.

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, *e.g.*, Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual*, Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (herein "Sambrook *et al.*, 1989"); *DNA Cloning: A Practical Approach*, Volumes I and II (D.N. Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed. 1984); *Nucleic Acid Hybridization* [B.D. Hames & S.J. Higgins eds. (1985)]; *Transcription And Translation* [B.D. Hames & S.J. Higgins, eds. (1984)]; *Animal Cell Culture* [R.I. Freshney, ed. (1986)]; *Immobilized Cells And Enzymes* [IRL Press, (1986)]; B. Perbal, *A Practical Guide To Molecular Cloning* (1984); F.M. Ausubel *et al.* (eds.), *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994).

Therefore, if appearing herein, the following terms shall have the definitions set out below.

As used herein, the term "gene" refers to an assembly of nucleotides that encode a polypeptide, and includes cDNA and genomic DNA nucleic acids.

30

A "vector" is a replicon, such as plasmid, phage or cosmid, to which another DNA segment may be attached so as to bring about the replication of the attached segment. A "replicon" is any genetic element (*e.g.*, plasmid, chromosome, virus) that functions

as an autonomous unit of DNA replication *in vivo*, *i.e.*, capable of replication under its own control.

5 A "cassette" refers to a segment of DNA that can be inserted into a vector at specific restriction sites. The segment of DNA encodes a polypeptide of interest, and the cassette and restriction sites are designed to ensure insertion of the cassette in the proper reading frame for transcription and translation.

10 A cell has been "transfected" by exogenous or heterologous DNA when such DNA has been introduced inside the cell.

A "nucleic acid molecule" refers to the phosphate ester polymeric form of ribonucleosides (adenosine, guanosine, uridine or cytidine; "RNA molecules") or deoxyribonucleosides (deoxyadenosine, deoxyguanosine, deoxythymidine, or deoxycytidine; "DNA molecules"), or any phosphoester analogues thereof, such as phosphorothioates and thioesters, in either single stranded form, or a double-stranded helix. Double stranded DNA-DNA, DNA-RNA and RNA-RNA helices are possible. The term nucleic acid molecule, and in particular DNA or RNA molecule, refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary forms. Thus, this term includes double-stranded DNA found, *inter alia*, in linear or circular DNA molecules (*e.g.*, restriction fragments), plasmids, and chromosomes. In discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the nontranscribed strand of DNA (*i.e.*, the strand having a sequence homologous to the mRNA). A "recombinant DNA molecule" is a DNA molecule that has undergone a molecular biological manipulation.

20  
25

A nucleic acid molecule is "hybridizable" to another nucleic acid molecule, such as a cDNA, genomic DNA, or RNA, when a single stranded form of the nucleic acid molecule can anneal to the other nucleic acid molecule under the appropriate conditions of temperature and solution ionic strength (*see Sambrook et al., supra*). The conditions of temperature and ionic strength determine the "stringency" of the

30

hybridization. For preliminary screening for homologous nucleic acids, low stringency hybridization conditions, corresponding to a  $T_m$  of  $55^\circ$ , can be used, *e.g.*, 5x SSC, 0.1% SDS, 0.25% milk, and no formamide; or 30% formamide, 5x SSC, 0.5% SDS). Moderate stringency hybridization conditions correspond to a higher  $T_m$ , *e.g.*, 40% formamide, with 5x or 6x SCC. High stringency hybridization conditions correspond to the highest  $T_m$ , *e.g.*, 50% formamide, 5x or 6x SCC. Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of  $T_m$  for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher  $T_m$ ) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating  $T_m$  have been derived (*see* Sambrook *et al.*, *supra*, 9.50-10.51). For hybridization with shorter nucleic acids, *i.e.*, oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (*see* Sambrook *et al.*, *supra*, 11.7-11.8). Preferably a minimum length for a hybridizable nucleic acid is at least about 12 nucleotides; preferably at least about 18 nucleotides; and more preferably the length is at least about 27 nucleotides; and most preferably 36 nucleotides.

In a specific embodiment, the term "standard hybridization conditions" refers to a  $T_m$  of  $55^\circ\text{C}$ , and utilizes conditions as set forth above. In a preferred embodiment, the  $T_m$  is  $60^\circ\text{C}$ ; in a more preferred embodiment, the  $T_m$  is  $65^\circ\text{C}$ .

A DNA "coding sequence" is a double-stranded DNA sequence which is transcribed and translated into a polypeptide in a cell *in vitro* or *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxyl) terminus. A coding sequence can include, but is not limited to, prokaryotic sequences and synthetic DNA sequences. If the coding sequence is



intended for expression in a eukaryotic cell, a polyadenylation signal and transcription termination sequence will usually be located 3' to the coding sequence.

Transcriptional and translational control sequences are DNA regulatory sequences,  
5 such as promoters, enhancers, terminators, and the like, that provide for the expression of a coding sequence in a host cell. In eukaryotic cells, polyadenylation signals are control sequences.

A "promoter sequence" is a DNA regulatory region capable of binding RNA  
10 polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence  
15 will be found a transcription initiation site (conveniently defined for example, by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase.

A coding sequence is "under the control" of transcriptional and translational control  
20 sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then trans-RNA spliced and translated into the protein encoded by the coding sequence.

A DNA sequence is "operatively linked" to an expression control sequence when the  
25 expression control sequence controls and regulates the transcription and translation of that DNA sequence. The term "operatively linked" includes having an appropriate start signal (e.g., ATG) in front of the DNA sequence to be expressed and maintaining the correct reading frame to permit expression of the DNA sequence under the control of the expression control sequence and production of the desired product encoded by  
30 the DNA sequence. If a gene that one desires to insert into a recombinant DNA molecule does not contain an appropriate start signal, such a start signal can be inserted in front of the gene.

As used herein, the term "homologous" in all its grammatical forms refers to the relationship between proteins that possess a "common evolutionary origin," including proteins from superfamilies (*e.g.*, the immunoglobulin superfamily) and homologous proteins from different species (*e.g.*, myosin light chain, etc.) [Reeck *et al.*, *Cell*, 5 50:667 (1987)]. Such proteins have sequence homology as reflected by their high degree of sequence similarity.

Accordingly, the term "sequence similarity" in all its grammatical forms refers to the degree of identity or correspondence between nucleic acid or amino acid sequences of 10 proteins that may or may not share a common evolutionary origin (*see* Reeck *et al.*, *supra*). However, in common usage and in the instant application, the term "homologous," when modified with an adverb such as "highly," may refer to sequence similarity and not a common evolutionary origin.

15 Two DNA sequences are "substantially homologous" when at least about 60% (preferably at least about 80%, and most preferably at least about 90 or 95%) of the nucleotides match over the defined length of the DNA sequences. Sequences that are substantially homologous can be identified by comparing the sequences using standard software available in sequence data banks, or in a Southern hybridization experiment 20 under, for example, stringent conditions as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, *e.g.*, Maniatis *et al.*, *supra*; DNA Cloning, Vols. I & II, *supra*; Nucleic Acid Hybridization, *supra*.

As used herein an amino acid sequence is 100% "homologous" to a second amino acid 25 sequence if the two amino acid sequences are identical, and/or differ only by neutral or conservative substitutions as defined below. Accordingly, an amino acid sequence is 50% "homologous" to a second amino acid sequence if 50% of the two amino acid sequences are identical, and/or differ only by neutral or conservative substitutions.

30 As used herein, DNA and protein sequence percent identity can be determined using MacVector 6.0.1, Oxford Molecular Group PLC (1996) and the Clustal W algorithm with the alignment default parameters, and default parameters for identity. These

commercially available programs can also be used to determine sequence similarity using the same or analogous default parameters.

5 The term "corresponding to" is used herein to refer similar or homologous sequences, whether the exact position is identical or different from the molecule to which the similarity or homology is measured. Thus, the term "corresponding to" refers to the sequence similarity, and not the numbering of the amino acid residues or nucleotide bases.

10 As used herein a "heterologous nucleotide sequence" is a nucleotide sequence that is added to a nucleotide sequence of the present invention by recombinant methods to form a nucleic acid which is not naturally formed in nature. Such nucleic acids can encode fusion proteins or peptides, including chimeric proteins and peptides. Thus the heterologous nucleotide sequence can encode peptides and/or proteins which contain  
15 regulatory and/or structural properties. In another such embodiment the heterologous nucleotide can encode a protein or peptide that functions as a means of detecting the protein or peptide encoded by the nucleotide sequence of the present invention after the recombinant nucleic acid is expressed. In still another such embodiment the heterologous nucleotide can function as a means of detecting a nucleotide sequence of  
20 the present invention. A heterologous nucleotide sequence can comprise non-coding sequences including restriction sites, regulatory sites, promoters and the like.

The present invention also relates to cloning vectors containing nucleic acids encoding analogs and derivatives of the bromodomains of the present invention and  
25 polypeptides/peptides that can bind a bromodomain when a lysine of the polypeptide/peptide is acetylated, including modified fragments, that have the same or homologous functional activity as the individual fragments, and homologs thereof. The production and use of derivatives and analogs related to the fragments are within the scope of the present invention.

30

Due to the degeneracy of nucleotide coding sequences, other DNA sequences which encode substantially the same amino acid sequence as a nucleic acid encoding a protein

comprising bromodomain or bromodomain binding partner (*i.e.*, when post-transcriptionally acetylated) of the present invention for example, may be used in the practice of the present invention. These include but are not limited to allelic genes, homologous genes from other species, which are altered by the substitution of different

5 codons that encode the same amino acid residue within the sequence, thus producing a silent change. Likewise, the peptides and polypeptides of the present invention include, but are not limited to, those containing, as a primary amino acid sequence, analogous portions of their respective amino acid sequences including altered sequences in which functionally equivalent amino acid residues are substituted for

10 residues within the sequence resulting in a conservative amino acid substitution. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity, which acts as a functional equivalent, resulting in a silent alteration. Substitutes for an amino acid within the sequence may be selected from other members of the class to which the amino acid belongs. For

15 example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. Amino acids containing aromatic ring structures are phenylalanine, tryptophan, and tyrosine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, and lysine.

20 The negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

Particularly preferred conserved amino acid exchanges are:

- (a) Lys for Arg or vice versa such that a positive charge may be maintained;
- (b) Glu for Asp or vice versa such that a negative charge may be maintained;
- 25 (c) Ser for Thr or vice versa such that a free -OH can be maintained;
- (d) Gln for Asn or vice versa such that a free NH<sub>2</sub> can be maintained;
- (e) Ile for Leu or for Val or vice versa as roughly equivalent hydrophobic amino acids; and
- (f) Phe for Tyr or vice versa as roughly equivalent aromatic amino acids.

30

A conservative change generally leads to less change in the structure and function of the resulting protein. A non-conservative change is more likely to alter the structure,

activity or function of the resulting protein. The present invention should be considered to include sequences containing conservative changes which do not significantly alter the activity or binding characteristics of the resulting protein. Specific amino acid residues for the P/CAF bromodomain have been identified that are  
5 important for binding, indicating a potential lower stringency for the substitution of the remaining amino acids residues.

All of the peptides/fragments of the present invention can be modified by being placed in a fusion or chimeric peptide or protein, or labeled *e.g.*, to have an N-terminal FLAG-  
10 tag, or H6 tag. In a particular embodiment the P/CAF bromodomain fragment can be modified to contain a marker protein such as green fluorescent protein as described in U.S. Patent No. 5,625,048 filed April 29, 1997 and WO 97/26333, published July 24, 1997 each of which are hereby incorporated by reference herein in their entireties.

15 The nucleic acids encoding peptides and protein fragments of the present invention and analogs thereof can be produced by various methods known in the art. The manipulations which result in their production can occur at the gene or protein level [Sambrook *et al.*, 1989, *supra*]. The nucleotide sequence can be cleaved at appropriate sites with restriction endonuclease(s), followed by further enzymatic modification if  
20 desired, isolated, and ligated *in vitro*. In addition a nucleic acid sequence can be mutated *in vitro* or *in vivo*, to create and/or destroy translation, initiation, and/or termination sequences, or to create variations in coding regions and/or form new restriction endonuclease sites or destroy preexisting ones, to facilitate further *in vitro* modification. Any technique for mutagenesis known in the art can be used, including  
25 but not limited to, *in vitro* site-directed mutagenesis [Hutchinson *et al.*, *J. Biol. Chem.*, 253:6551 (1978); Zoller and Smith, *DNA*, 3:479-488 (1984); Oliphant *et al.*, *Gene*, 44:177 (1986); Hutchinson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 83:710 (1986)], use of TAB® linkers (Pharmacia), etc. PCR techniques are preferred for site directed mutagenesis [see Higuchi, 1989, "Using PCR to Engineer DNA", in *PCR Technology: Principles and Applications for DNA Amplification*, H. Erlich, ed., Stockton Press,  
30 Chapter 6, pp. 61-70].

The identified and isolated nucleic acids can then be inserted into an appropriate cloning vector. A large number of vector-host systems known in the art may be used.

### Protein expression and purification

5

A bacterial protein expression system can be used to make various stable isotopically labeled ( $^{13}\text{C}$ ,  $^{15}\text{N}$ , and  $^2\text{H}$ ) protein samples that are useful for a three-dimensional NMR structural determination of a protein complex. For example a pET14b (Novagen) bacterial expression vector can be constructed which expresses the recombinant P/CAF  
10 bromodomain as an amino-terminal His-tagged fusion protein.

Protein expression and purification can be conducted using standard procedures for His-tagged proteins [Zhou *et al.*, *J. Biol. Chem.* **270**:31119-31123 (1995)]. To optimize the level of protein expression, various bacterial growth and expression  
15 conditions can be screened, which include different *E. Coli* cell lines, and growth and protein induction temperatures. Generally, it is preferred to obtain the maximum amount of soluble protein while still inducing protein expression with a relatively low IPTG concentration *e.g.*,  $\sim 0.2\text{mM}$  (final concentration) at  $16^\circ\text{C}$ . As exemplified below, the bromodomain of P/CAF (residues 719-832 of SEQ ID NO:2 which is SEQ  
20 ID NO:7) was subcloned into the pET14b expression vector (Novagen) and expressed in *Escherichia coli* BL21(DE3) cells. Uniformly  $^{15}\text{N}$ - and  $^{15}\text{N}/^{13}\text{C}$ -labeled proteins were prepared by growing bacteria in a minimal medium containing  $^{15}\text{NH}_4\text{Cl}$  with or without  $^{13}\text{C}_6$ -glucose. A uniformly  $^{15}\text{N}/^{13}\text{C}$ -labeled and fractionally deuterated protein sample was prepared by growing the cells in 75%  $^2\text{H}_2\text{O}$ . The bromodomain was  
25 purified by affinity chromatography on a nickel-IDA column (Invitrogen) followed by the removal of poly-His tag by thrombin cleavage. The final purification of the protein was achieved by size-exclusion chromatography. The acetyl-lysine-containing peptides were prepared on a MilliGen 9050 peptide synthesizer (Perkin Elmer) using Fmoc/HBTU chemistry. Acetyl-lysine was incorporated using the reagent  
30 Fmoc-Ac-Lys with HBTU/DIPEA activation. NMR samples contained approximately 1 mM protein in 100mM phosphate buffer of pH 6.5 and 5mM perdeuterated DTT and 0.5mM EDTA in  $\text{H}_2\text{O}/^2\text{H}_2\text{O}$  (9/1) or  $^2\text{H}_2\text{O}$ .

One major advantage of using the heteronuclear multidimensional approach, as exemplified herein, is that the NMR resonance assignments of a protein are obtained in a sequence-specific manner which assures accuracy and greatly facilitates data analysis and structure determination [Clore, G. M. & Gronenborn, A. M. *Meth. Enzymol.* 239:249-363 (1994)]. In addition, the signal overlapping problems in the protein spectra are minimized by the use of multidimensional NMR spectra, which separates the proton signals according to the chemical shifts of their attached hetero-nuclei (such as  $^{15}\text{N}$  and  $^{13}\text{C}$ ). This NMR approach has been proven very powerful for structural analysis of large proteins [Clore, G. M. & Gronenborn, A. M. *Meth. Enzymol.* 239:249-363 (1994)]. To facilitate sequence-specific resonance assignments for the structural study, a uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled and fractionally (75%) deuterated protein sample of the bromodomain can be prepared by growing bacterial cells in 75%  $^2\text{H}_2\text{O}$  as exemplified below. Such protein samples can be used for triple-resonance NMR experiments. A triple-labeled protein sample is useful for high-resolution NMR structural studies. Because of the favorable  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  relaxation rates caused by the partial deuteration of the protein, constant-time triple-resonance NMR spectra can be acquired with higher digital resolution and sensitivity [Sattler, M. & Fesik, S. W. *Structure* 4:1245-1249 (1996)]. In addition, various stable-isotopically labeled ( $^{15}\text{N}$  and  $^{13}\text{C}$  /  $^{15}\text{N}$ ) proteins can also be prepared using this procedure.

20

### Synthetic Polypeptides

The term "polypeptide" is used in its broadest sense to refer to a compound of two or more subunit amino acids, amino acid analogs, or peptidomimetics. The subunits are linked by peptide bonds. The terms "polypeptide", "protein", and "peptide" are used interchangeably herein, though preferably as used herein a "peptide" refers to a compound of at least two but less than fifty subunit amino acids, and a polypeptide or protein refers to compound of fifty or more amino acids. The polypeptides of the present invention may be chemically synthesized or as detailed above, genetically engineered or isolated from natural sources.

30

In addition, potential drugs or agents that may be tested in the drug screening assays of the present invention may also be chemically synthesized. When the peptide is to be

modified, *e.g.*, acetylated, the modification can be at any time during the peptide synthesis, including using an acetyl-lysine as a starting material or acetylating a lysine residue of a peptide after the peptide has been synthesized. In the Example below, the acetyl-lysine-containing peptides were prepared on a MilliGen 9050 peptide synthesizer (Perkin Elmer) using Fmoc/HBTU chemistry. Acetyl-lysine was incorporated using the reagent Fmoc-Ac-Lys with HBTU/DIPEA activation.

Thus, synthetic polypeptides, prepared using the well known techniques of solid phase, liquid phase, or peptide condensation techniques, or any combination thereof, can include natural and unnatural amino acids. Amino acids used for peptide synthesis may be standard Boc ( $N^{\alpha}$ -amino protected  $N^{\alpha}$ -t-butyloxycarbonyl) amino acid resin with the standard deprotecting, neutralization, coupling and wash protocols of the original solid phase procedure of Merrifield [*J. Am. Chem. Soc.*, **85**:2149-2154 (1963)], or the base-labile  $N^{\alpha}$ -amino protected 9-fluorenylmethoxycarbonyl (Fmoc) amino acids first described by Carpino and Han [*J. Org. Chem.*, **37**:3403-3409 (1972)]. Both Fmoc and Boc  $N^{\alpha}$ -amino protected amino acids can be obtained from Fluka, Bachem, Advanced Chemtech, Sigma, Cambridge Research Biochemical, Bachem, or Peninsula Labs or other chemical companies familiar to those who practice this art. In addition, the method of the invention can be used with other  $N^{\alpha}$ -protecting groups that are familiar to those skilled in this art. Solid phase peptide synthesis may be accomplished by techniques familiar to those in the art and provided, for example, in Stewart and Young [Solid Phase Synthesis, Second Edition, Pierce Chemical Co., Rockford, IL (1984)] and Fields and Noble [*Int. J. Pept. Protein Res.*, **35**:161-214 (1990)], or using automated synthesizers, such as sold by ABS. Thus, polypeptides of the invention may comprise D-amino acids, a combination of D- and L-amino acids, and various "designer" amino acids (*e.g.*,  $\beta$ -methyl amino acids,  $C\alpha$ -methyl amino acids, and  $N\alpha$ -methyl amino acids, etc.) to convey special properties. Synthetic amino acids include ornithine for lysine, fluorophenylalanine for phenylalanine, and norleucine for leucine or isoleucine. Additionally, by assigning specific amino acids at specific coupling steps,  $\alpha$ -helices,  $\beta$  turns,  $\beta$  sheets,  $\gamma$ -turns, and cyclic peptides can be generated.



In a further embodiment, subunits of peptides that confer useful chemical and structural properties will be chosen. For example, peptides comprising D-amino acids will be resistant to L-amino acid-specific proteases *in vivo*. In addition, the present invention envisions preparing peptides that have more well defined structural

5 properties, and the use of peptidomimetics, and peptidomimetic bonds, such as ester bonds, to prepare peptides with novel properties. In another embodiment, a peptide may be generated that incorporates a reduced peptide bond, i.e.,  $R_1\text{-CH}_2\text{-NH-R}_2$ , where  $R_1$  and  $R_2$  are amino acid residues or sequences. A reduced peptide bond may be introduced as a dipeptide subunit. Such a molecule would be resistant to peptide bond

10 hydrolysis, *e.g.*, protease activity. Such peptides would provide ligands with unique function and activity, such as extended half-lives *in vivo* due to resistance to metabolic breakdown, or protease activity. Furthermore, it is well known that in certain systems constrained peptides show enhanced functional activity [Hruby, *Life Sciences*, 31:189-199 (1982); Hruby *et al.*, *Biochem J.*, 268:249-262 (1990)]; the present invention

15 provides a method to produce a constrained peptide that incorporates random sequences at all other positions.

*Constrained and cyclic peptides.* A constrained, cyclic or rigidized peptide may be prepared synthetically, provided that in at least two positions in the sequence of the

20 peptide an amino acid or amino acid analog is inserted that provides a chemical functional group capable of crosslinking to constrain, cyclise or rigidize the peptide after treatment to form the crosslink. Cyclization will be favored when a turn-inducing amino acid is incorporated. Examples of amino acids capable of crosslinking a peptide are cysteine to form disulfides, aspartic acid to form a lactone or a lactam, and a

25 chelator such as  $\gamma$ -carboxyl-glutamic acid (Gla) (Bachem) to chelate a transition metal and form a cross-link. Protected  $\gamma$ -carboxyl glutamic acid may be prepared by modifying the synthesis described by Zee-Cheng and Olson [*Biophys. Biochem. Res. Commun.*, 94:1128-1132 (1980)]. A peptide in which the peptide sequence comprises at least two amino acids capable of crosslinking may be treated, *e.g.*, by oxidation of

30 cysteine residues to form a disulfide or addition of a metal ion to form a chelate, so as to crosslink the peptide and form a constrained, cyclic or rigidized peptide.

The present invention provides strategies to systematically prepare cross-links. For example, if four cysteine residues are incorporated in the peptide sequence, different protecting groups may be used (Hiskey, in *The Peptides: Analysis, Synthesis, Biology*, Vol. 3, Gross and Meienhofer, eds., Academic Press: New York, pp. 137-167 (1981);  
 5 Ponsanti *et al.*, *Tetrahedron*, **46**:8255-8266 (1990)]. The first pair of cysteines may be deprotected and oxidized, then the second set may be deprotected and oxidized. In this way a defined set of disulfide cross-links may be formed. Alternatively, a pair of cysteines and a pair of chelating amino acid analogs may be incorporated so that the cross-links are of a different chemical nature.

10

*Non-classical amino acids that induce conformational constraints.* The following non-classical amino acids may be incorporated in the peptide in order to introduce particular conformational motifs: 1,2,3,4-tetrahydroisoquinoline-3-carboxylate [Kazmierski *et al.*, *J. Am. Chem. Soc.*, **113**:2275-2283 (1991)]; (2S,3S)-methyl-  
 15 phenylalanine, (2S,3R)-methyl-phenylalanine, (2R,3S)-methyl-phenylalanine and (2R,3R)-methyl-phenylalanine (Kazmierski and Hruby, *Tetrahedron Lett.* (1991)]; 2-aminotetrahydronaphthalene-2-carboxylic acid [Landis, Ph.D. Thesis, University of Arizona (1989)]; hydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylate [Miyake *et al.*, *J. Takeda Res. Labs.*, **43**:53-76 (1989)];  $\beta$ -carboline (D and L) [Kazmierski, Ph.D.  
 20 Thesis, University of Arizona (1988)]; HIC (histidine isoquinoline carboxylic acid) [Zechel *et al.*, *Int. J. Pep. Protein Res.*, **43** (1991)]; and HIC (histidine cyclic urea) (Dharanipragada).

The following amino acid analogs and peptidomimetics may be incorporated into a  
 25 peptide to induce or favor specific secondary structures: LL-Acp (LL-3-amino-2-propenidone-6-carboxylic acid), a  $\beta$ -turn inducing dipeptide analog [Kemp *et al.*, *J. Org. Chem.*, **50**:5834-5838 (1985)];  $\beta$ -sheet inducing analogs [Kemp *et al.*, *Tetrahedron Lett.*, **29**:5081-5082 (1988)];  $\beta$ -turn inducing analogs [Kemp *et al.*, *Tetrahedron Lett.*, **29**:5057-5060 (1988)];  $\alpha$ -helix inducing analogs (Kemp *et al.*,  
 30 *Tetrahedron Lett.*, **29**:4935-4938 (1988)];  $\gamma$ -turn inducing analogs [Kemp *et al.*, *J. Org. Chem.*, **54**:109:115 (1989)]; and analogs provided by the following references: Nagai and Sato, *Tetrahedron Lett.*, **26**:647-650 (1985); DiMaio *et al.*, *J. Chem. Soc. Perkin Trans.*, p. 1687 (1989); also a Gly-Ala turn analog [Kahn *et al.*, *Tetrahedron*

spectra of mutated proteins can be compared to that of the wild-type protein bromodomain.

Chemical-shift perturbations due to ligand binding have proven to be a reliable and sensitive probe for the ligand binding site of the protein. This is because the chemical-shift changes of the backbone amide groups are likely to reflect any changes in protein conformation and/or hydrogen bonding due to the peptide/ligand binding. To examine the effects of a mutation on the ligand binding (in this case the ligand is a peptide comprising an acetyl-lysine), peptide titration experiments can be conducted by following the changes of  $^1\text{H}/^{15}\text{N}$  signals of the mutant proteins as a function of the peptide concentration. These experiments indicate whether the acetyl-lysine binding site remains the same or changes in the mutants relative to the wild type protein. The effects of the mutation on the peptide binding affinity can also be examined by NMR spectroscopy. If the mutated proteins result in the reduction of the binding affinity, a change of the exchange phenomenon between the free and the ligand-bound signals should be observed in NMR spectrum. If the reduction in binding affinity causes the peptide binding to change from a slow exchange rate to a fast exchange rate, on the NMR time scale, then the peptide binding affinity can be determined from the NMR titration experiment. From these mutation analyses key amino acid residues that are important for binding a peptide comprising the acetyl-lysine can be identified. Such analysis has been exemplified below.

#### Protein Structure Determination by NMR Spectroscopy

The NMR results from the present invention are summarized by the atomic structure coordinates of the free form of the P/CAF bromodomain (Table 5) and of the P/CAF bromodomain-acetyl-histamine complex (Table 6). The NMR chemical shift assignments of the P/CAF bromodomain are included in the chemical shift table (Table 1) for the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of P/CAF bromodomain. The unambiguous NOE-derived Inter-proton Distance Restraints are in Table 2, the ambiguous NOE-derived Inter-proton Distance Restraints are in Table 3, and the  $^1\text{H}$  bonding restraints are disclosed in Table 4.

*Backbone and Side-chain Assignments:* Sequence-specific backbone assignment can be achieved by using a suite of deuterium-decoupled triple-resonance 3D NMR experiments which include HNCA, HN(CO)CA, HN(CA)CB, HN(COCA)CB, HNCO, and HN(CA)CO experiments [Yamazaki, *et al.*, *J. Am. Chem. Soc.* **116**:11655-11666 (1994)]. The water flip-back scheme is used in these NMR pulse programs to minimize amide signal attenuation from water exchange. Sequential side-chain assignments are typically accomplished from a series of 3D NMR experiments with alternative approaches to confirm the assignments. These experiments include 3D  $^{15}\text{N}$  TOCSY-HSQC, HCCH-TOCSY, (H)C(CO)NH-TOCSY, and H(C)(CO)NH-TOCSY [see Clore, G. M. & Gronenborn, A. M. *Meth. Enzymol.* **239**:249-363 (1994); Sattler *et al.*, *Prog. in Nuclear Magnetic Resonance Spec.* **4**:93-158 (1999)].

*Stereospecific Methyl Groups:* Stereospecific assignments of methyl groups of Valine and Leucine residues can be obtained from an analysis of carbon signal multiplet splitting using a fractionally  $^{13}\text{C}$ -labeled protein sample, which can be readily prepared using M9 minimal medium containing 10%  $^{13}\text{C}$ -/90%  $^{12}\text{C}$ -glucose mixture [see Neri, *et al.*, *Biochemistry* **28**:7510-7516 (1989)].

*Dihedral Angle Restraints:* Backbone dihedral angle ( $\Phi$ ) constraints can be generated from the  $^3J_{\text{HNH}\alpha}$  coupling constants measured in a HNHA-*J* experiment [see Vuister, G. & Bax, A. *J. Am. Chem. Soc.* **115**:7772-7777 (1993)]. Side-chain dihedral angles ( $\chi_1$ ) can be obtained from short mixing time  $^{15}\text{N}$ -edited 3D TOCSY-HSQC [see Clore, *et al.*, *J. Biomol. NMR* **1**:13-22 (1991)] and 3D HNHB experiments [see Matson *et al.*, *J. Biomol. NMR* **3**:239-244 (1993)], which can also provide stereospecific assignments of  $\beta$  methylene protons.

*Hydrogen Bonds Restraints:* Amide protons that are involved in hydrogen bonds can be identified from an analysis of amide exchange rates measured from a series of 2D  $^1\text{H}/^{15}\text{N}$  HSQC spectra recorded after adding  $^2\text{H}_2\text{O}$  to the protein sample.

*NOE Distance Restraints:* Distance restraints are obtained from analysis of  $^{15}\text{N}$ , and  $^{13}\text{C}$ -edited 3D NOESY data, which can be collected with different mixing times to minimize spin diffusion problems. The nuclear Overhauser effect (NOE)-derived

restraints are categorized as strong (1.8-3 Å), medium (1.8-4 Å) or weak (1.8-5 Å) based on the observed NOE intensities. A recently developed procedure for the iterative automated NOE analysis by using ARIA [see Nilges *et al.*, *Prog. NMR Spectroscopy* 32:107-139 (1998)] can be employed which integrates with X-PLOR for structural calculations. To ensure the success of ARIA/X-PLOR-assisted NOE analysis and structure calculations, the ARIA assigned NOE peaks can be manually confirmed.

*Intermolecular NOE Distance Restraints:* For the structural determination of a protein/peptide complex, intermolecular NOE distance restraints can be obtained from a  $^{13}\text{C}$ -edited ( $F_1$ ) and  $^{15}\text{N}$ , and  $^{13}\text{C}$ -filtered ( $F_3$ ) 3D NOESY data set collected for a sample containing isotope-labeled protein and non-labeled peptide.

*Structure Calculations and Refinements:* Structures of the protein can be generated using a distance geometry/simulated annealing protocol with the X-PLOR program [see Nilges, *et al.*, *FEBS Lett.* 229:317-324 (1988); Kuszewski, *et al.*, *J. Biomol. NMR* 2:33-56 (1992); Brünger, A. T. *X-PLOR Version 3.1: A system for X-Ray crystallography and NMR* (Yale University Press, New Haven, CT, 1993)]. The structure calculations can employ inter-proton distance restraints obtained from  $^{15}\text{N}$ - and  $^{13}\text{C}$ -resolved NOESY spectra. The initial low-resolution structures can be used to facilitate NOE assignments, and help identify hydrogen bonding partners for slowly exchanging amide protons. The experimental restraints of dihedral angles and hydrogen bonds can be included in the distance restraints for structure refinements.

#### Protein-Structure Based Design of Agonists and Antagonists of the Bromodomain-Acetyl-Lysine Binding Complex

Once the three-dimensional structure of the Bromodomain and the Bromodomain-acetyl-lysine binding complex are determined, a potential drug or agent (antagonist or agonist) can be examined through the use of computer modeling using a docking program such as GRAM, DOCK, or AUTODOCK [Dunbrack *et al.*, 1997, *supra*]. This procedure can include computer fitting of potential agents to the bromodomain, for example, to ascertain how well the shape and the chemical structure of the potential ligand will complement or interfere with the interaction between the bromodomain and

the acetyl-lysine [Bugg *et al.*, *Scientific American*, **Dec.**:92-98 (1993); West *et al.*, *TIPS*, **16**:67-74 (1995)]. Computer programs can also be employed to estimate the attraction, repulsion, and steric hindrance of the agent to the dimer-dimer binding site, for example. Generally the tighter the fit (*e.g.*, the lower the steric hindrance, and/or  
5 the greater the attractive force) the more potent the potential drug will be since these properties are consistent with a tighter binding constant. Furthermore, the more specificity in the design of a potential drug the more likely that the drug will not interfere with related proteins. This will minimize potential side-effects due to unwanted interactions with other proteins.

10

Initially a potential drug could be obtained by screening a random peptide library produced by recombinant bacteriophage for example, [Scott and Smith, *Science*, **249**:386-390 (1990); Cwirla *et al.*, *Proc. Natl. Acad. Sci.*, **87**:6378-6382 (1990); Devlin *et al.*, *Science*, **249**:404-406 (1990)] or a chemical library. An agent selected in  
15 this manner could be then be systematically modified by computer modeling programs until one or more promising potential drugs are identified. Such analysis has been shown to be effective in the development of HIV protease inhibitors [Lam *et al.*, *Science* **263**:380-384 (1994); Wlodawer *et al.*, *Ann. Rev. Biochem.* **62**:543-585 (1993); Appelt, *Perspectives in Drug Discovery and Design* **1**:23-48 (1993); Erickson,  
20 *Perspectives in Drug Discovery and Design* **1**:109-128 (1993)].

Such computer modeling allows the selection of a finite number of rational chemical modifications, as opposed to the countless number of essentially random chemical modifications that could be made, any one of which might lead to a useful drug. Each  
25 chemical modification requires additional chemical steps, which while being reasonable for the synthesis of a finite number of compounds, quickly becomes overwhelming if all possible modifications needed to be synthesized. Thus, through the use of the three-dimensional structural analysis disclosed herein and computer modeling, a large number of these compounds can be rapidly screened on the computer  
30 monitor screen, and a few likely candidates can be determined without the laborious synthesis of untold numbers of compounds.

Once a potential drug (agonist or antagonist) is identified it can be either selected from a library of chemicals as are commercially available from most large chemical companies including Merck, GlaxoWellcome, Bristol Meyers Squib, Monsanto/Searle, Eli Lilly, Novartis and Pharmacia UpJohn, or alternatively the potential drug may be synthesized *de novo*. As mentioned above, the *de novo* synthesis of one or even a relatively small group of specific compounds is reasonable in the art of drug design.

The potential drug can then be tested in any standard binding assay (including in high throughput binding assays) for its ability to bind to the ZA loop of a bromodomain. Alternatively the potential drug can be tested for its ability to modulate the binding of a bromodomain to acetylated histamine, for example. When a suitable potential drug is identified, a second NMR structural analysis can optionally be performed on the binding complex formed between the bromodomain-acetyl-lysine binding complex, or the bromodomain alone and the potential drug. Computer programs that can be used to aid in solving such three-dimensional structures include QUANTA, CHARMM, INSIGHT, SYBYL, MACROMODE, and ICM, MOLMOL, RASMOL, AND GRASP [Kraulis, *J. Appl Crystallogr.* 24:946-950 (1991)]. Most if not all of these programs and others as well can be also obtained from the WorldWideWeb through the internet. Using the approach described herein and equipped with the structural analysis disclosed herein, the three-dimensional structures of other bromodomain-acetyl-lysine binding complexes can more readily be obtained and analyzed. Such analysis will, in turn, allow corresponding drug screening methodology to be performed using the three-dimensional structures of such related complexes.

For all of the drug screening assays described herein further refinements to the structure of the drug will generally be necessary and can be made by the successive iterations of any and/or all of the steps provided by the particular drug screening assay, including further structural analysis by NMR, for example.

30

#### *Phage libraries for Drug Screening.*

Phage libraries have been constructed which when infected into host *E. coli* produce random peptide sequences of approximately 10 to 15 amino acids [Parmley and Smith,

Gene 73:305-318 (1988), Scott and Smith, Science 249:386-249 (1990)]. Specifically, the phage library can be mixed in low dilutions with permissive *E. coli* in low melting point LB agar which is then poured on top of LB agar plates. After incubating the plates at 37°C for a period of time, small clear plaques in a lawn of *E. coli* will form which represents active phage growth and lysis of the *E. coli*. A representative of these phages can be absorbed to nylon filters by placing dry filters onto the agar plates. The filters can be marked for orientation, removed, and placed in washing solutions to block any remaining absorbent sites. The filters can then be placed in a solution containing, for example, a radioactive bromodomain. After a specified incubation period, the filters can be thoroughly washed and developed for autoradiography. Plaques containing the phage that bind to the radioactive bromodomain can then be identified. These phages can be further cloned and then retested for their ability to bind to the bromodomain as before. Once the phage has been purified, the binding sequence contained within the phage can be determined by standard DNA sequencing techniques. Once the DNA sequence is known, synthetic peptides can be generated which are encoded by these sequences. These peptides can be tested, for example, for their ability to modulate the affinity of the bromodomain for its binding partner (*e.g.*, a protein comprising an acetyl-lysine or a fragment of that protein).

The effective peptide(s) can be synthesized in large quantities for use in *in vivo* models and eventually in humans to treat certain tumors. It should be emphasized that synthetic peptide production is relatively non-labor intensive, easily manufactured, quality controlled and thus, large quantities of the desired product can be produced quite cheaply. Similar combinations of mass produced synthetic peptides have been used with great success [Patarroyo, *Vaccine*, 10:175-178 (1990)].

#### Drug Screening Assays

The drug screening assays of the present invention may use any of a number of means for determining the interaction between an agent or drug and a peptide comprising an acetyl-lysine and/or a bromodomain. Thus, standard high throughput drug screening procedures can be employed using a library of low molecular weight compounds, for



example that can be screened to identify a binding partner for the bromodomain. Any such chemical library can be used including those discussed above.

In a particular assay, a bromodomain is placed on or coated onto a solid support.

5 Methods for placing the peptides or proteins on the solid support are well known in the art and include such things as linking biotin to the protein and linking avidin to the solid support. An agent is allowed to equilibrate with the bromodomain to test for binding. Generally, the solid support is washed and agents that are retained are selected as potential drugs. Alternatively, a peptide comprising an acetyl-lysine is  
10 placed on or coated onto a solid support. In a particular embodiment of this type, the peptide comprises the amino acid sequence of SEQ ID NO:4.

The agent may be labeled. For example, in one embodiment radiolabeled agents are used to measure the binding of the agent. In another embodiment the agents have  
15 fluorescent markers. In yet another embodiment, a Biocore chip (Pharmacia) coated with the bromodomain is used, for example and the change in surface conductivity can be measured.

In addition, since a number of proteins have been identified that contain  
20 bromodomains, and the binding partners of many of these proteins are known, the fact that the bromodomain specifically binds to an acetylated lysine as disclosed herein allows the identification and preparation of a number of potential modulators of the bromodomain-acetyl-lysine binding complex based on the amino acid sequences of the binding partners to the proteins. Such potential modulators include : ISYGR-AcK-  
25 KRRQRR (SEQ ID NO:4), ARKSTGG-AcK-APRKQL (SEQ ID NO:5) and QSTSRHK-AcK-LMFKTE (SEQ ID NO:6) which bind to the P/CAF bromodomain as shown in the Example, below. Such peptides also can be used, for example, as a starting point for the design of an inhibitor of the bromodomain-acetyl-lysine binding complex.

30

Alternatively, a drug can be specifically designed to bind to the ZA loop of a bromodomain for example, such as the P/CAF bromodomain, and be assayed through NMR based methodology [Shuker *et al.*, *Science* 274:1531-1534 (1996) hereby

incorporated by reference in its entirety.] In a particular embodiment, analogs of the binding partner of the bromodomain can be used in this analysis. One such peptide has the amino acid sequence of SEQ ID NO:4. In another embodiment of this type, the peptide has the amino acid sequence of SEQ ID NO:5. In another such embodiment of  
5 this type, the peptide has the amino acid sequence of SEQ ID NO:6.

The assay begins with contacting a compound with a  $^{15}\text{N}$ -labeled bromodomain. Binding of the compound with the ZA loop of the bromodomain can be determined by monitoring the  $^{15}\text{N}$ - or  $^1\text{H}$ -amide chemical shift changes in two dimensional  $^{15}\text{N}$ -  
10 heteronuclear single-quantum correlation ( $^{15}\text{N}$ -HSQC) spectra upon the addition of the compound to the  $^{15}\text{N}$ -labeled bromodomain. Since these spectra can be rapidly obtained, it is feasible to screen a large number of compounds [Shuker *et al.*, *Science* 274:1531-1534 (1996)]. A compound is identified as a potential ligand if it binds to the ZA loop of the bromodomain. In a further embodiment, the potential ligand can  
15 then be used as a model structure, and analogs to the compound can be obtained (e.g, from the vast chemical libraries commercially available, or alternatively through *de novo* synthesis). The analogs are then screened for their ability to bind the ZA loop of the bromodomain thus to obtain a ligand. An analog of the potential ligand is chosen as a ligand when it binds to the ZA loop of the bromodomain with a higher binding  
20 affinity than the potential ligand. In a preferred embodiment of this type the analogs are screened by monitoring the  $^{15}\text{N}$ - or  $^1\text{H}$ -amide chemical shift changes in two dimensional  $^{15}\text{N}$ -heteronuclear single-quantum correlation ( $^{15}\text{N}$ -HSQC) spectra upon the addition of the analog to the  $^{15}\text{N}$ -labeled bromodomain as described above.

25 In another further embodiment, compounds are screened for binding to two nearby sites on the bromodomain. In this case, a compound that binds a first site of the bromodomain does not bind a second nearby site. Binding to the second site can be determined by monitoring changes in a different set of amide chemical shifts in either the original screen or a second screen conducted in the presence of a ligand (or  
30 potential ligand) for the first site. From an analysis of the chemical shift changes the approximate location of a potential ligand for the second site is identified. Optimization of the second ligand for binding to the site is then carried out by screening structurally related compounds (e.g., analogs as described above). When

- ligands for the first site and the second site are identified, their location and orientation in the ternary complex can be determined experimentally either by NMR spectroscopy or X-ray crystallography. On the basis of this structural information, a linked compound is synthesized in which the ligand for the first site and the ligand for the
- 5 second site are linked. In a preferred embodiment of this type the two ligands are covalently linked. This linked compound is tested to determine if it has a higher binding affinity for the bromodomain than either of the two individual ligands. A linked compound is selected as a ligand when it has a higher binding affinity for the bromodomain than either of the two ligands. In a preferred embodiment the affinity of
- 10 the linked compound with the bromodomain is determined monitoring the  $^{15}\text{N}$ - or  $^1\text{H}$ -amide chemical shift changes in two dimensional  $^{15}\text{N}$ -heteronuclear single-quantum correlation ( $^{15}\text{N}$ -HSQC) spectra upon the addition of the linked compound to the  $^{15}\text{N}$ -labeled bromodomain as described above.
- 15 A larger linked compound can be constructed in an analogous manner, *e.g.*, linking three ligands which bind to three nearby sites on the bromodomain to form a multilinked compound that has an even higher affinity for the bromodomain than the linked compound.

20

#### Identification of New Bromodomains

- By disclosing that protein bound acetyl-lysine is a binding partner for bromodomains, the present invention provides a method of identifying novel proteins that contain bromodomains. In short, a protein fragment or analog thereof comprising an acetyl-
- 25 lysine can be used as bait to identify a binding partner that comprises a bromodomain. Any one of a number of procedures can be carried out to identify such a binding partner. One such assay comprises passing a cell extract over the bait peptide which is attached to a solid support. After washing the solid support to remove any non-specific binders, the bromodomain containing protein can be eluted from the solid
- 30 support with an appropriate eluant. In a particular embodiment, the free bait peptide can be used in the elution. Other methodology includes the use of a yeast two-hybrid system, a GST pull down assay, ELISA, immunometric assays, and a modification of the CORT procedure of Schlessinger *et al.*, (US Patent No. 5,858,686, Issued on

January 12, 1999 which is hereby incorporated by reference in its entirety) for use with the bromodomain-acetyl-lysine binding complex.

#### Labels:

5

Suitable labels include enzymes, fluorophores (*e.g.*, fluorescein isothiocyanate (FITC), phycoerythrin (PE), Texas red (TR), rhodamine, free or chelated lanthanide series salts, especially  $\text{Eu}^{3+}$ , to name a few fluorophores), chromophores, radioisotopes, chelating agents, dyes, colloidal gold, latex particles, ligands (*e.g.*, biotin), and  
 10 chemiluminescent agents. When a control marker is employed, the same or different labels may be used for the test and control marker gene.

In the instance where a radioactive label, such as the isotopes  $^3\text{H}$ ,  $^{14}\text{C}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^{36}\text{Cl}$ ,  $^{51}\text{Cr}$ ,  $^{57}\text{Co}$ ,  $^{58}\text{Co}$ ,  $^{59}\text{Fe}$ ,  $^{90}\text{Y}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ , and  $^{186}\text{Re}$  are used, known currently available  
 15 counting procedures may be utilized. In the instance where the label is an enzyme, detection may be accomplished by any of the presently utilized colorimetric, spectrophotometric, fluorospectrophotometric, amperometric or gasometric techniques known in the art.

20 Direct labels are one example of labels which can be used according to the present invention. A direct label has been defined as an entity, which in its natural state, is readily visible, either to the naked eye, or with the aid of an optical filter and/or applied stimulation, *e.g.* U.V. light to promote fluorescence. Among examples of colored labels, which can be used according to the present invention, include metallic sol  
 25 particles, for example, gold sol particles such as those described by Leuving (U.S. Patent 4,313,734); dye sol particles such as described by Gribnau *et al.* (U.S. Patent 4,373,932 and May *et al.* (WO 88/08534); dyed latex such as described by May, *supra*, Snyder (EP-A 0 280 559 and 0 281 327); or dyes encapsulated in liposomes as described by Campbell *et al.* (U.S. Patent 4,703,017). Other direct labels include a  
 30 radionucleotide, a fluorescent moiety or a luminescent moiety. In addition to these direct labeling devices, indirect labels comprising enzymes can also be used according to the present invention. Various types of enzyme linked immunoassays are well known in the art, for example, alkaline phosphatase and horseradish peroxidase,

lysozyme, glucose-6-phosphate dehydrogenase, lactate dehydrogenase, urease, these and others have been discussed in detail by Eva Engvall in Enzyme Immunoassay ELISA and EMIT in *Methods in Enzymology*, 70:419-439 (1980) and in U.S. Patent 4,857,453.

5

Suitable enzymes include, but are not limited to, alkaline phosphatase,  $\beta$ -galactosidase, green fluorescent protein and its derivatives, luciferase, and horseradish peroxidase.

Other labels for use in the invention include magnetic beads or magnetic resonance  
10 imaging labels.

#### Antibodies to Portions of the Bromodomain that Interact with Acetyl-Lysine

According to the present invention, the bromodomains, and more particularly the ZA  
15 loops of the bromodomains and fragments thereof can be produced by a recombinant source, or through chemical synthesis, or through the modification of these peptides and fragments; and derivatives or analogs thereof, including fusion proteins, may be used as an immunogen to generate antibodies that specifically interfere with the formation of the bromodomain-acetyl-lysine binding complex. Similarly, antibodies  
20 can be raised against peptides that comprise one or more acetyl-lysine residues which also interfere with the formation of the bromodomain-acetyl-lysine binding complex. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments, and a Fab expression library.

25 Various procedures known in the art may be used for the production of the polyclonal antibodies. For the production of antibody, various host animals can be immunized by injection with the peptide having the amino acid sequence of SEQ ID NO:3, for example, or a derivative (*e.g.*, or fusion protein) thereof, including but not limited to rabbits, mice, rats, sheep, goats, etc. In one embodiment, the peptide can be  
30 conjugated to an immunogenic carrier, *e.g.*, bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH). Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface

active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (*bacille Calmette-Guerin*) and *Corynebacterium parvum*.

- 5 For preparation of monoclonal antibodies directed toward the peptides or protein fragments of the present invention, or analog, or derivative thereof, any technique that provides for the production of antibody molecules by continuous cell lines in culture may be used. These include but are not limited to the hybridoma technique originally developed by Kohler and Milstein [*Nature*, **256**:495-497 (1975)], as well as the trioma  
10 technique, the human B-cell hybridoma technique [Kozbor *et al.*, *Immunology Today*, **4**:72 (1983); Cote *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, **80**:2026-2030 (1983)], and the EBV-hybridoma technique to produce human monoclonal antibodies [Cole *et al.*, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96 (1985)]. In an additional embodiment of the invention, monoclonal antibodies can be produced in  
15 germ-free animals utilizing technology described in PCT/US90/02545. In fact, according to the invention, techniques developed for the production of "chimeric antibodies" [Morrison *et al.*, *J. Bacteriol.*, **159**:870 (1984); Neuberger *et al.*, *Nature*, **312**:604-608 (1984); Takeda *et al.*, *Nature*, **314**:452-454 (1985)] by splicing the genes from a mouse antibody molecule specific for the peptide having the amino acid  
20 sequence of SEQ ID NO:3, for example, together with genes from a human antibody molecule of appropriate biological activity can be used; such antibodies are within the scope of this invention. Such human or humanized chimeric antibodies are preferred for use in therapy of human diseases or disorders (described *infra*), since the human or humanized antibodies are much less likely than xenogenic antibodies to induce an  
25 immune response, in particular an allergic response, themselves.

According to the invention, techniques described for the production of single chain antibodies [U.S. Patent Nos. 5,476,786 and 5,132,405 to Huston; U.S. Patent 4,946,778] can be adapted to produce specific single chain antibodies. An additional  
30 embodiment of the invention utilizes the techniques described for the construction of Fab expression libraries [Huse *et al.*, *Science*, **246**:1275-1281 (1989)] to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

Antibody fragments which contain the idiotype of the antibody molecule can be generated by known techniques. For example, such fragments include but are not limited to: the  $F(ab')_2$  fragment which can be produced by pepsin digestion of the antibody molecule; the Fab' fragments which can be generated by reducing the  
 5 disulfide bridges of the  $F(ab')_2$  fragment, and the Fab fragments which can be generated by treating the antibody molecule with papain and a reducing agent.

In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, *e.g.*, radioimmunoassay, ELISA (enzyme-linked  
 10 immunosorbant assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitin reactions, immunodiffusion assays, *in situ* immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), western blots, precipitation reactions, agglutination assays (*e.g.*, gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays,  
 15 and immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labeled. Many means are known in the art for detecting binding in an immunoassay and are  
 20 within the scope of the present invention. For example, to select antibodies which recognize a specific epitope of a ZA loop of a bromodomain, for example, one may assay generated hybridomas for a product which binds to a bromodomain fragment containing such an epitope and choose those which do not cross-react with bromodomain fragments that do not include that epitope.

25

In a specific embodiment, antibodies that interfere with the formation of the bromodomain-acetyl-lysine complex can be generated. Such antibodies can be tested using the assays described and could potentially be used in anti-cancer therapies.

30

#### Administration

According to the invention, the component or components of a therapeutic composition, *e.g.*, an agent of the invention that interferes with the bromodomain-

acetyl-lysine binding complex such as the peptide having the amino acid sequence of SEQ ID NOs:4, 5, or 6 and a pharmaceutically acceptable carrier, may be introduced parenterally, transmucosally, *e.g.*, orally, nasally, or rectally, or transdermally.

Preferably, administration is parenteral, *e.g.*, via intravenous injection, and also including, but is not limited to, intra-arteriole, intramuscular, intradermal, subcutaneous, intraperitoneal, intraventricular, and intracranial administration.

In a preferred aspect, the agent of the present invention can cross cellular and nuclear membranes, which would allow for intravenous or oral administration. Strategies are available for such crossing, including but not limited to, increasing the hydrophobic nature of a molecule; introducing the molecule as a conjugate to a carrier, such as a ligand to a specific receptor, targeted to a receptor; and the like.

The present invention also provides for conjugating targeting molecules to such an agent. "Targeting molecule" as used herein shall mean a molecule which, when administered *in vivo*, localizes to desired location(s). In various embodiments, the targeting molecule can be a peptide or protein, antibody, lectin, carbohydrate, or steroid. In one embodiment, the targeting molecule is a peptide ligand of a receptor on the target cell. In a specific embodiment, the targeting molecule is an antibody. Preferably, the targeting molecule is a monoclonal antibody. In one embodiment, to facilitate crosslinking the antibody can be reduced to two heavy and light chain heterodimers, or the F(ab')<sub>2</sub> fragment can be reduced, and crosslinked to the agent via the reduced sulfhydryl. Antibodies for use as targeting molecule are specific for a cell surface antigen.

In another embodiment, the therapeutic compound can be delivered in a vesicle, in particular a liposome [see Langer, *Science*, **249**:1527-1533 (1990); Treat *et al.*, in *Liposomes in the Therapy of Infectious Disease and Cancer*, Lopez-Berestein and Fidler (eds.), Liss: New York, pp. 353-365 (1989); Lopez-Berestein, *ibid.*, pp. 317-327; see generally *ibid.*].

In yet another embodiment, the therapeutic compound can be delivered in a controlled release system. For example, the agent may be administered using intravenous



infusion, an implantable osmotic pump, a transdermal patch, liposomes, or other modes of administration. In one embodiment, a pump may be used [see Langer, *supra*; Sefton, *CRC Crit. Ref. Biomed. Eng.*, **14**:201 (1987); Buchwald *et al.*, *Surgery*, **88**:507 (1980); Saudek *et al.*, *N. Engl. J. Med.*, **321**:574 (1989)]. In another embodiment, 5 polymeric materials can be used [see *Medical Applications of Controlled Release*, Langer and Wise (eds.), CRC Press: Boca Raton, Florida (1974); *Controlled Drug Bioavailability, Drug Product Design and Performance*, Smolen and Ball (eds.), Wiley: New York (1984); Ranger and Peppas, *J. Macromol. Sci. Rev. Macromol. Chem.*, **23**:61 (1983); see also Levy *et al.*, *Science*, **228**:190 (1985); During *et al.*, *Ann.* 10 *Neurol.*, **25**:351 (1989); Howard *et al.*, *J. Neurosurg.*, **71**:105 (1989)]. In yet another embodiment, a controlled release system can be placed in proximity of the therapeutic target, *i.e.*, the bone marrow, thus requiring only a fraction of the systemic dose [see, *e.g.*, Goodson, in *Medical Applications of Controlled Release, supra*, vol. 2, pp. 115-138 (1984)]. Other controlled release systems are discussed in the review by Langer 15 [*Science*, **249**:1527-1533 (1990)].

*Pharmaceutical Compositions.* In yet another aspect of the present invention, provided are pharmaceutical compositions of the above. Such pharmaceutical compositions may be for administration for injection, or for oral, pulmonary, nasal or other forms of 20 administration. In general, comprehended by the invention are pharmaceutical compositions comprising effective amounts of a low molecular weight component or components, or derivative products, of the invention together with pharmaceutically acceptable diluents, preservatives, solubilizers, emulsifiers, adjuvants and/or carriers. Such compositions include diluents of various buffer content (*e.g.*, Tris-HCl, acetate, 25 phosphate), pH and ionic strength; additives such as detergents and solubilizing agents (*e.g.*, Tween 80, Polysorbate 80), anti-oxidants (*e.g.*, ascorbic acid, sodium metabisulfite), preservatives (*e.g.*, Thimersol, benzyl alcohol) and bulking substances (*e.g.*, lactose, mannitol); incorporation of the material into particulate preparations of polymeric compounds such as polylactic acid, polyglycolic acid, etc. or into liposomes. 30 Hylauronic acid may also be used. Such compositions may influence the physical state, stability, rate of *in vivo* release, and rate of *in vivo* clearance of the present proteins and derivatives. See, *e.g.*, Remington's Pharmaceutical Sciences, 18th Ed. [1990, Mack Publishing Co., Easton, PA 18042] pages 1435-1712 which are herein

incorporated by reference. The compositions may be prepared in liquid form, or may be in dried powder, such as lyophilized form.

*Oral Delivery.* Contemplated for use herein are oral solid dosage forms, which are described generally in Remington's Pharmaceutical Sciences, 18th Ed. 1990 (Mack Publishing Co. Easton PA 18042) at Chapter 89, which is herein incorporated by reference. Solid dosage forms include tablets, capsules, pills, troches or lozenges, cachets or pellets. Also, liposomal or proteinoid encapsulation may be used to formulate the present compositions (as, for example, proteinoid microspheres reported in U.S. Patent No. 4,925,673). Liposomal encapsulation may be used and the liposomes may be derivatized with various polymers (e.g., U.S. Patent No. 5,013,556). A description of possible solid dosage forms for the therapeutic is given by Marshall, K. In: *Modern Pharmaceutics* Edited by G.S. Banker and C.T. Rhodes Chapter 10, 1979, herein incorporated by reference. In general, the formulation will include an agent of the present invention (or chemically modified forms thereof) and inert ingredients which allow for protection against the stomach environment, and release of the biologically active material in the intestine.

Also specifically contemplated are oral dosage forms of the above derivatized component or components. The component or components may be chemically modified so that oral delivery of the derivative is efficacious. Generally, the chemical modification contemplated is the attachment of at least one moiety to the component molecule itself, where said moiety permits (a) inhibition of proteolysis; and (b) uptake into the blood stream from the stomach or intestine. Also desired is the increase in overall stability of the component or components and increase in circulation time in the body. An example of such a moiety is polyethylene glycol.

For the component (or derivative) the location of release may be the stomach, the small intestine (the duodenum, the jejunum, or the ileum), or the large intestine. One skilled in the art has available formulations which will not dissolve in the stomach, yet will release the material in the duodenum or elsewhere in the intestine. Preferably, the release will avoid the deleterious effects of the stomach environment, either by

protection of the protein (or derivative) or by release of the biologically active material beyond the stomach environment, such as in the intestine.

5 The therapeutic can be included in the formulation as fine multi-particulates in the form of granules or pellets of particle size about 1 mm. The formulation of the material for capsule administration could also be as a powder, lightly compressed plugs or even as tablets. The therapeutic could be prepared by compression.

10 One may dilute or increase the volume of the therapeutic with an inert material. These diluents could include carbohydrates, especially mannitol,  $\alpha$ -lactose, anhydrous lactose, cellulose, sucrose, modified dextrans and starch. Certain inorganic salts may be also be used as fillers including calcium triphosphate, magnesium carbonate and sodium chloride. Some commercially available diluents are Fast-Flo, Emdex, STA-Rx 1500, Emcompress and Avicell.

15 Disintegrants may be included in the formulation of the therapeutic into a solid dosage form. Materials used as disintegrates include but are not limited to starch, including the commercial disintegrant based on starch, Explotab. Binders also may be used to hold the therapeutic agent together to form a hard tablet and include materials from  
20 natural products such as acacia, tragacanth, starch and gelatin.

An anti-frictional agent may be included in the formulation of the therapeutic to prevent sticking during the formulation process. Lubricants may be used as a layer between the therapeutic and the die wall. Glidants that might improve the flow  
25 properties of the drug during formulation and to aid rearrangement during compression also might be added. The glidants may include starch, talc, pyrogenic silica and hydrated silicoaluminate.

In addition, to aid dissolution of the therapeutic into the aqueous environment a  
30 surfactant might be added as a wetting agent. Additives which potentially enhance uptake of the protein (or derivative) are for instance the fatty acids oleic acid, linoleic acid and linolenic acid.

*Nasal Delivery.* Nasal delivery of an agent of the present invention (or derivative) is also contemplated. Nasal delivery allows the passage of a peptide, for example, to the blood stream directly after administering the therapeutic product to the nose, without the necessity for deposition of the product in the lung. Formulations for nasal delivery  
 5 include those with dextran or cyclodextran.

*Transdermal administration.* Various and numerous methods are known in the art for transdermal administration of a drug, *e.g.*, via a transdermal patch. Transdermal patches are described in for example, U.S. Patent No. 5,407,713, issued April 18, 1995  
 10 to Rolando *et al.*; U.S. Patent No. 5,352,456, issued October 4, 1994 to Fallon *et al.*; U.S. Patent No. 5,332,213 issued August 9, 1994 to D'Angelo *et al.*; U.S. Patent No. 5,336,168, issued August 9, 1994 to Sibalís; U.S. Patent No. 5,290,561, issued March 1, 1994 to Farhadieh *et al.*; U.S. Patent No. 5,254,346, issued October 19, 1993 to Tucker *et al.*; U.S. Patent No. 5,164,189, issued November 17, 1992 to Berger *et al.*;  
 15 U.S. Patent No. 5,163,899, issued November 17, 1992 to Sibalís; U.S. Patent Nos. 5,088,977 and 5,087,240, both issued February 18, 1992 to Sibalís; U.S. Patent No. 5,008,110, issued April 16, 1991 to Benecke *et al.*; and U.S. Patent No. 4,921,475, issued May 1, 1990 to Sibalís, the disclosure of each of which is incorporated herein by reference in its entirety.

20

It can be readily appreciated that a transdermal route of administration may be enhanced by use of a dermal penetration enhancer, *e.g.*, such as enhancers described in U.S. Patent No. 5,164,189 (*supra*), U.S. Patent No. 5,008,110 (*supra*), and U.S. Patent No. 4,879,119, issued November 7, 1989 to Aruga *et al.*, the disclosure of each of  
 25 which is incorporated herein by reference in its entirety.

*Pulmonary Delivery.* Also contemplated herein is pulmonary delivery of the pharmaceutical compositions of the present invention. A pharmaceutical composition of the present invention is delivered to the lungs of a mammal while inhaling and  
 30 traverses across the lung epithelial lining to the blood stream. Other reports of this include Adjei *et al.* [*Pharmaceutical Research*, 7:565-569 (1990); Adjei *et al.*, *International Journal of Pharmaceutics*, 63:135-144 (1990) (leuprolide acetate); Braquet *et al.*, *Journal of Cardiovascular Pharmacology*, 13(suppl. 5):143-146 (1989)

(endothelin-1); Hubbard *et al.*, *Annals of Internal Medicine*, Vol. III, pp. 206-212 (1989) ( $\alpha$ 1-antitrypsin); Smith *et al.*, *J. Clin. Invest.*, 84:1145-1146 (1989) ( $\alpha$ -1-proteinase); Oswein *et al.*, "Aerosolization of Proteins", *Proceedings of Symposium on Respiratory Drug Delivery II*, Keystone, Colorado, March, (1990) (recombinant human growth hormone); Debs *et al.*, *J. Immunol.*, 140:3482-3488 (1988) (interferon- $\gamma$  and tumor necrosis factor alpha); Platz *et al.*, U.S. Patent No. 5,284,656 (granulocyte colony stimulating factor)]. A method and composition for pulmonary delivery of drugs for systemic effect is described in U.S. Patent No. 5,451,569, issued September 19, 1995 to Wong *et al.*

10

A subject in whom administration of an agent of the present invention is an effective therapeutic regiment for cancer, for example, is preferably a human, but can be any animal. Thus, as can be readily appreciated by one of ordinary skill in the art, the methods and pharmaceutical compositions of the present invention are particularly suited to administration to any animal, *e.g.*, for veterinary medical use, particularly for a mammal, and including, but by no means limited to, domestic animals, such as feline or canine subjects, farm animals, including bovine, equine, caprine, ovine, and porcine subjects, wild animals (whether in the wild or in a zoological garden), research animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats, avian species, such as chickens, turkeys, and songbirds.

20

The present invention may be better understood by reference to the following non-limiting Example, which is provided as exemplary of the invention. The following example is presented in order to more fully illustrate the preferred embodiments of the invention. It should in no way be construed, however, as limiting the broad scope of the invention.

25

30

## 5

10 Indeed, although this motif has also been identified in other chromatin proteins,  
heretofore not even one binding partner for a bromodomain had been identified.

**Sample preparation:** The bromodomain of P/CAF (residues 719-832 of SEQ ID NO:2)

was subcloned into the pET14b expression vector (Novagen) and expressed in *Escherichia coli* BL21(DE3) cells. Uniformly  $^{15}\text{N}$ - and  $^{15}\text{N}/^{13}\text{C}$ -labelled proteins were prepared by growing bacteria in a minimal medium containing  $^{15}\text{NH}_4\text{Cl}$  with or without  $^{13}\text{C}_6$ -glucose. A uniformly  $^{15}\text{N}/^{13}\text{C}$ -labelled and fractionally deuterated protein sample was prepared by growing the cells in 75%  $^2\text{H}_2\text{O}$ . The bromodomain was purified by affinity chromatography on a nickel-IDA column (Invitrogen) followed by the removal of poly-His tag by thrombin cleavage. The final purification of the protein was achieved by size-exclusion chromatography. The acetyl-lysine-containing peptides were prepared on a MilliGen 9050 peptide synthesizer (Perkin Elmer) using Fmoc/HBTU chemistry. Acetyl-lysine was incorporated using the reagent Fmoc-Ac-Lys with HBTU/DIPEA activation. NMR samples contained approximately 1 mM protein in 100mM phosphate buffer of pH 6.5 and 5mM perdeuterated DTT and 0.5mM EDTA in  $\text{H}_2\text{O}/^2\text{H}_2\text{O}$  (9/1) or  $^2\text{H}_2\text{O}$ .

*NMR spectroscopy:* All NMR spectra were acquired at 30°C on a Bruker DRX600 or DRX500 spectrometer. The backbone assignments of the  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  resonances were achieved using deuterium-decoupled triple-resonance experiments of HNCACB and HN(CO)CACB [Yamazaki *et al.*, *J. Am. Chem. Soc.* **116**:11655-11666 (1994)] recorded using the uniformly  $^{15}\text{N}/^{13}\text{C}$ -labeled and fractionally deuterated protein. The

side-chain atoms were assigned from 3D HCCH-TOCSY [Clare and Gronenborn, *Meth. Enzymol.* **239**:249-363 (1994)] and (H)C(CO)NH-TOCSY [Logan *et al.*, *J. Biomol. NMR* **3**:225-231 (1993)] data collected on the uniformly  $^{15}\text{N}/^{13}\text{C}$ -labeled protein. Stereospecific assignments of methyl groups of the Val and Leu residues were  
 5 obtained using a fractionally  $^{13}\text{C}$ -labeled sample [Neri *et al.*, *Biochemistry* **28**:7510-7516 (1989)]. The NOE-derived distance restraints were obtained from  $^{15}\text{N}$ - or  $^{13}\text{C}$ -edited 3D NOESY spectra.  $\phi$ -angle restraints were determined based on the  $^3J_{\text{HN,H}\alpha}$  coupling constants measured in a 3D HNHA spectrum [Clare and Gronenborn, *Meth. Enzymol.* **239**:249-363 (1994)]. Slowly exchanging amide protons were  
 10 identified from a series of 2D  $^{15}\text{N}$ -HSQC spectra recorded after the  $\text{H}_2\text{O}$  buffer was changed to a  $^2\text{H}_2\text{O}$  buffer. The intermolecular NOEs used in defining the structure of the bromodomain/Ac-histamine complex were detected in  $^{13}\text{C}$ -edited ( $F_1$ ),  $^{13}\text{C}/^{15}\text{N}$ -filtered ( $F_2$ ) 3D NOESY spectrum [Clare and Gronenborn, *Meth. Enzymol.* **239**:249-363 (1994)]. All NMR spectra were processed with the NMRPipe/NMRDraw  
 15 programs and analyzed using NMRView [Johnson and Blevins, *J. Biomol. NMR* **4**:603-614 (1994)].

*Structure calculations:* Structures of the bromodomain were calculated with a distance geometry/simulated annealing protocol using the X-PLOR program [Brunger, A. *X-PLOR Version 3.1: A system for X-Ray crystallography and NMR*, Yale University  
 20 Press, New Haven, CT, (1993)]. A total of 1324 manually assigned NOE-derived distance restraints were obtained from the  $^{15}\text{N}$ - and  $^{13}\text{C}$ -edited NOE spectra. Further analysis of the NOE spectra was carried out by the iterative automated assignment procedure using ARIA [Nilges and O'Donoghue, *Prog. NMR Spectroscopy* **32**:107-139  
 25 (1998)], which integrates with X-PLOR for structure calculations. A total of 1519 unambiguous and 590 ambiguous distance restraints were identified from the NOE data by ARIA, many of which were checked and confirmed manually. The ARIA-assigned distance restraints were in agreement with the structures calculated using only the manually assigned NOE distance restraints, 28 hydrogen-bond distance  
 30 restraints for 14 hydrogen bonds, and 54  $\phi$ -angle restraints. The final structure calculations employed a total of 3515 NMR experimental restraints obtained from the manual and the ARIA-assisted assignments, 2843 of which were unambiguously assigned NOE-derived distance restraints that comprise of 1077 intra-residue, 621

sequential, 550 medium-range, and 595 long-range NOEs. For the ensemble of the final 30 structures, no distance and torsional angle restraints were violated by more than 0.3 Å and 5°, respectively. The total, distance violation, and dihedral violation energies were  $178.7 \pm 2.4$  kcal mol<sup>-1</sup>,  $41.6 \pm 0.9$  kcal mol<sup>-1</sup>, and  $0.50 \pm 0.06$  kcal mol<sup>-1</sup>, respectively. The Lennard-Jones potential which was not used during any refinement stage, was  $-526.2 \pm 16.8$  kcal mol<sup>-1</sup> for the final structures. Ramachandran plot analysis of the final structures (residues 727-828) with Procheck-NMR [Laskowski *et al.*, *J. Biomol. NMR* 8:477-486 (1996)] showed that  $71.0 \pm 0.6\%$ ,  $23.8 \pm 0.6\%$ ,  $3.5 \pm 0.2\%$ , and  $1.7 \pm 0.2\%$  of the non-Gly and non-Pro residues were in the most favorable, additionally allowed, generously allowed, and disallowed regions, respectively. The corresponding values for the residues in the four  $\alpha$ -helices (residues 727-743, 770-776, 785-802, and 807-827) were  $88.9 \pm 0.4\%$ ,  $11.0 \pm 0.4\%$ ,  $0.1 \pm 0.1\%$ , and  $0.0 \pm 0.0\%$ , respectively. The structure of the bromodomain/acetyl-histamine complex was determined using the free form structure and additional 25 intermolecular and 5 intra-ligand NOE-derived distance restraints.

*Site-directed mutagenesis:* Mutant proteins were prepared using the QuickChange site-directed mutagenesis kit (Stratagene). The presence of appropriate mutations was confirmed by DNA sequencing.

20

*Ligand titration:* Ligand titration experiments were performed by recording a series of 2D <sup>15</sup>N- and <sup>13</sup>C-HSQC spectra on the uniformly <sup>15</sup>N-, and <sup>15</sup>N/<sup>13</sup>C-labelled bromodomain (~0.3mM), respectively, in the presence of different amounts of ligand concentration ranging from 0 to approximately 2.0 mM. The protein sample and the stock solutions of the ligands were all prepared in the same aqueous buffer containing 100mM phosphate and 5mM perdeuterated DTT at pH 6.5.

25

The full length nucleic acid sequence of the human p300/CBP-associated factor (P/CAF) was obtained from GenBank. Accession No: U57317.2 (SEQ ID NO:1) :

30

```

1  ggggccgcgt cgacgcggaa aagaggccgt ggggggcctc ccagcgctgg cagacaccgt
61  gaggctggca gccgccggca cgcacaccta gtccgcagtc ccgaggaaca tgtccgcagc
121  cagggcgcg ggcagagtc cgggcaggag aaccaaggga gggcggtgtgc tgtggcggcg
181  gcggcagcgg cagcggagcc gctagtcccc tccctcctgg gggagcagct gccgccgctg
241  ccgccgccgc caccaccatc agcgcgcggg gcccggccag agcgagccgg gcgagcggcg

```



301 cgctaggggg agggcggggg cggggagggg ggtgggcgaa gggggcgggg gggcgtgggg  
 361 ggaggggtctc gctctcccga ctaccagagc ccgagggaga ccctggcggc ggcggcgggc  
 421 cctgacactc ggcgccctct gccgtgctcc gggggcgcat gtccgaggct ggcggggccg  
 481 ggccgggagg ctgccccga ggagccgggg cagggggcgg gcccgggcg ctgcccccg  
 5 541 agcctgcggc gcttccgccc gcgccccgc agggctcccc ctgcgcccgt gccgccccgg  
 601 gctcgggcgc ctgcggtccg gcgacggcag tggctgcagc gggcacggcc gaaggaccgg  
 661 gaggcgggtg ctcgccccga atcgccgtga agaaagcgca actacgctcc gctccgcggg  
 721 ccaagaaact ggagaaactc ggagtgtact ccgcctgcaa ggccgaggag tcttgtaaat  
 781 gtaatggctg gaaaaaccct aaccctcac ccactcccc cagagccgac ctgcagcaaa  
 10 841 taattgtcag tctaacagaa tcctgtcgga gttgtagcca tgccctagct gctcatgttt  
 901 cccacctgga gaatgtgtca gaggaagaaa tgaacagact cctgggaata gtattggatg  
 961 tggaaatatct ctttacctgt gtccacaagg aagaagatgc agataccaaa caagtttatt  
 1021 tctatctatt taagctcttg agaaagtcta ttttaciaag agggaaacct gtgggtgaag  
 1081 gctcttttga aaagaaaccc ccatttga aaacctagcat tgaacagggt gtgaataact  
 15 1141 ttgtgcagta caaatttagt cacctgccag caaaagaaag gcaaacaata gttgagttgg  
 1201 caaaaatgtt cctaaaccgc atcaactatt ggcactctga ggcaccatct caacgaagac  
 1261 tgcgatctcc caatgatgat atttctggat acaaagagaa ctacacaagg tggctgtgtt  
 1321 actgcaacgt gccacagttc tgcgacagtc tacctcggtg cgaaaccaca cagggtgtttg  
 1381 ggagaacatt gcttcgctcg gtcttctact ttatgaggcg acaactcctg gaacaagcaa  
 20 1441 gacaggaaaa agataaaactg cctcttga aaacgaactct aatcctcact catttcccaa  
 1501 aatttctgtc catgctagaa gaagaagtat atagtcaaaa ctctcccatc tgggatcagg  
 1561 attttctctc agcctcttcc agaaccagcc agctaggcat ccaaacagtt atcaatccac  
 1621 ctctgtggc tgggacaatt tcatacaatt caacctcatc ttccttgag cagccaaacg  
 1681 caggagagcag cagtcctgcc tgcaaacgct cttctggact tgaggcaaac ccaggagaaa  
 25 1741 agaggaaaaat gactgattct catgttctgg aggaggcaa gaaaccccg gttatggggg  
 1801 atattccgat ggaattaatc aacgagggtta tgtctaccat cacggacct gcagcaatgc  
 1861 ttggaccaga gaccaatttt ctgtcagcac actcgccag ggatgaggcg gcaagggttg  
 1921 aagagcgagc ggggtgtaatt gaatttcacg tgggttgcaa ttcctcaac cagaaacca  
 1981 acaagaagat cctgatgtgg ctggttgccc tacagaacgt tttctccac cagctgcccc  
 30 2041 gaatgccaaa agaatacatc acacggctcg tctttgaccc gaaacacaaa acccttgctt  
 2101 taattaaaga tggccgtgtt attgggtgga tctgtttccg tatgttccca tctcaaggat  
 2161 tcacagagat tgtcttctgt gctgtaacct caaatgagca agtcaaggcg tatggaacac  
 2221 acctgatgaa tcatttga aaatatacaca taaagcatga catcctgaac ttcctcacat  
 2281 atgcagatga atatgcaatt ggatacttta agaaacaggg tttctccaaa gaaattaaaa  
 35 2341 tacctaaaac caaatatgtt ggctatatca aggattatga aggagccact ttaatgggat  
 2401 gtgagctaaa tccacggatc ccgtacacag aattttctgt catcattaaa aagcagaagg  
 2461 agataattaa aaaactgatt gaaagaaaac aggcacaaat tcgaaaagtt taccctggac  
 2521 tttcatgttt taaagatgga gttcgacaga ttcctataga aagcattcct ggaattagag  
 2581 agacaggctg gaaaccgagt ggaaaagaga aaagtaaaga gccagagac cctgaccagc  
 40 2641 ttacagcac gctcaagagc atcctccagc aggtgaagag ccatcaaagc gcttggccct  
 2701 tcatggaacc tgtgaagaga acagaagctc caggatatta tgaagttata aggttcccca  
 2761 tggatctgaa aaccatgagt gaacgcctca agaataaggta ctacgtgtct aagaaattat  
 2821 tcatggcaga cttacagcga gtctttacca attgcaaaga gtacaacgcc gctgagagt  
 2881 aatactacaa atgtgccaat atcctggaga aattcttctt cagtaaaatt aaggaagctg

2941 gattaattga caagtgattt tttttccccc tctgcttctt agaaactcac caagcagtgt  
3001 gcctaaagca aggt

The full length protein sequence of the human p300/CBP-associated factor (P/CAF)

5 was obtained from GenBank. Accession No: U57317.2, (SEQ ID NO:2):

1 MSEAGGAGPG GCGAGAGAGA GPGALPPQPA ALPPAPPQGS PCAAAGGSG ACPATAVAA  
61 AGTAEGPGGG GSARIAVKKA QLSAPRAKK LEKLGVSAC KAEESCKCNG WKNPNPSPTP  
121 PRADLQQIIV SLTESCRSCS HALAAHVSHL ENVSEEEMNR LLGIVLDVEY LFTCVHKEED  
181 ADTKQVYFYL FKLLRKSILQ RGKPVVEGSL EKKPPFEKPS IEQGVNMFVQ YKFSHLPAGE  
10 241 RQTIVELAKM FLNRINYWHL EAPSQRRLRS PNDDISGYKE NYTRWLCYCN VPQFCDLPR  
301 YETTQVFGRT LLRSVFTVMR RQLLEQARQE KDKLPLEKRT LILTHFPKFL SMLEEEVYSQ  
361 NSPIWDQDFL SASSRSQLG IQTVINPPPV AGTISYNSTS SSLEQPNAGS SSPACKASSG  
421 LEANPGEKRR MTDSHVLEEA KKPRVMGDIP MELINEVMST ITDPAAMLGP ETNFLSAHSA  
481 RDEAARLEER RGVIEFHVVG NSLNQKPNKK ILMWLVGLQN VFSHQLPRMP KEYITRLVFD  
15 541 PKHKTALAIK DGRVIGGICF RMFPSQGFTE IVFCAVTSNE QVKGYGTHLM NHLKEYHIKH  
601 DILNFLTAD EYAIGYFKKQ GFSKEIKIPK TKYVGYIKDY EGATLMGCEL NPRIPTYEFS  
661 VIKKQKEII KKLIERKQAQ IRKVYPGLSC FKDGVRQIPI ESIPGIRETG WKPSGKEKSK  
721 EPRDPDQLYS TLKSILQQVK SHQSAWPFME PVKRTAPGY YEVIRFPMDL KTMSERLKNR  
781 YYVSKKLFMA DLQRVFTNCK EYNAAESEYY KCANILEKFF FSKIKEAGLI DK

20

## Results

The P/CAF bromodomain represents an extensive family of bromodomains (Figure 1). A large number of long-range nuclear Overhauser enhancement (NOE)-derived  
25 distance restraints were identified in the NMR data of the P/CAF bromodomain, yielding a well-defined three-dimensional structure (Figures 2A -2D). Table 1 shows the NMR chemical shift assignment of the P/CAF bromodomain. Table 2 shows the Unambiguous NOE-derived distance restraints. Table 3 shows the Ambiguous NOE-derived distance restraints. Table 4 shows the Hydrogen bond restraints. The NMR  
30 structure coordinates of the P/CAF bromodomain in the free and complexed to acetyl-histamine are shown in Tables 5 and 6, respectively.

The structure consists of a four-helix bundle (helices  $\alpha_Z$ ,  $\alpha_A$ ,  $\alpha_B$ , and  $\alpha_C$ ) with a left-handed twist, and a long intervening loop between helices  $\alpha_Z$  and  $\alpha_A$  (termed the  
35 ZA loop, Figure 2E). The four amphipathic  $\alpha$ -helices are packed tightly against one another in an antiparallel manner, with crossing angles for adjacent helices of  $\sim 16$ - $20^\circ$ . The up-and-down four-helix bundle can adapt two topological folds with opposite

handedness (Figures 2F-2G). The right-handed four-helix bundle fold occurs more commonly and is seen in proteins such as hemerythrin and cytochrome  $b_{562}$ . The left-handed fold of the bromodomain structure is less common, but also observed in proteins such as cytochrome  $b_5$  and T4 lysozyme [Richardson, J., *Adv. Protein Chem.*, 34:167-339 (1989); Presnell and Cohen, *Proc. Natl. Acad. Sci. USA* 86:6592-6596 (1989)]. This topological difference arises from the orientation of the loop between the first two helices (Fig. 2F-2G). The right-handed four-helix bundle proteins have a relatively short hairpin-like connection between the first two helices, which makes the “preferred” turn to the right at the top of the first helix [Richardson, J., *Adv. Protein Chem.*, 34:167-339 (1989); Presnell and Cohen, *Proc. Natl. Acad. Sci. USA* 86:6592-6596 (1989); Weber and Salemme, *Nature* 287:82-84 (1980)]. In contrast, proteins with the left-handed fold usually have a long loop after the first helix and often contain additional secondary structural elements at the base of the helix bundle [Richardson, J., *Adv. Protein Chem.*, 34:167-339 (1989); Presnell and Cohen, *Proc. Natl. Acad. Sci. USA* 86:6592-6596 (1989)]. In the bromodomain structure, this long ZA loop has a defined conformation and is packed against the loop between helices  $\alpha_B$  and  $\alpha_C$  (termed the BC loop) to form a hydrophobic pocket. These tertiary interactions between the two loops appear to favor the left turn of the ZA loop, resulting in the left-handed four-helix bundle fold of the bromodomain. The hydrophobic pocket formed by loops ZA and BC is lined by residues Val752, Ala757, Tyr760, Val763, Tyr802 and Tyr809 (Fig. 2H), and appears to be a site for protein-protein interactions (see below). The pocket is located at one end of the four-helix bundle, opposite to the N- and C-termini of the protein. Interestingly, the ZA loop varies in length amongst different bromodomains, but almost always contains residues corresponding to Phe748, Pro751, Pro758, Tyr760, and Pro767 (Figure 1). The conservation of these residues within the ZA loop as well as residues within the  $\alpha$ -helical regions implies a similar left-handed four-helix bundle structure for the large family of bromodomains (Fig. 1).

The modular bromodomain structure supports the idea that bromodomain can act as a functional unit for protein-protein interactions. The observation that bromodomains are found in nearly all known nuclear HATs (A-type) that are known to promote transcription-related acetylation of histones on specific lysine residues, but not present in cytoplasmic HATs (B-type), prompted the determination of whether bromodomains

can interact with acetyl-lysine (AcK). The NMR titration of the P/CAF bromodomain were performed with a peptide (SGRGKGG-AcK-GLGK) derived from histone H4, in which Lys8 is acetylated (Lys8 is the major acetylation site in H4 for GCN5, a yeast homologue of P/CAF). Remarkably, the bromodomain could indeed bind the AcK peptide. Moreover, this interaction appeared to be specific, based on the  $^{15}\text{N}$ -HSQC spectra which showed that only a limited number of residues underwent chemical shift changes as a function of peptide concentration (Figure 3A). Conversely, the NMR titration of the bromodomain with a non-acetylated, but otherwise identical H4 peptide, showed no noticeable chemical shift changes, demonstrating that the interaction between the bromodomain and the lysine-acetylated H4 peptide was dependent upon acetylation of lysine. The dissociation constant ( $K_D$ ) for the AcK peptide was estimated to be  $346 \pm 54 \mu\text{M}$ . This binding is likely reinforced through additional interactions between bromodomain-containing proteins and target proteins. Notably, many chromatin-associated proteins contain two or multiple bromodomains (Figure 1). Indeed, binding with another lysine-acetylated peptide (RKSTGG-AcK-APRKQ) derived from the major acetylation site on histone H3 (residues 9-20) was also observed. Together, these data demonstrate that the P/CAF bromodomain has the ability to bind AcK peptides in an acetylation dependent manner.

Intriguingly, the bromodomain residues that exhibited the most significant  $^1\text{H}$  and  $^{15}\text{N}$  chemical shift changes on peptide binding are located near the hydrophobic pocket between the ZA and BC loops (Figure 3B). Because a similar pattern of amide chemical shift changes was observed with the two different AcK-containing peptides, it was surmised that the hydrophobic cavity is the primary binding site for AcK. This hypothesis was further supported by titration with acetyl-histamine, which mimics the chemical structure of the AcK side-chain (Figure 3C). Both  $^{15}\text{N}$ - and  $^{13}\text{C}$ -HSQC spectra showed that interaction with acetyl-histamine was also acetylation-dependent, involving the same set of residues that showed chemical shift perturbations with similar concentration dependence. It should be noted that the bromodomain did not bind to the amino acids acetyl-lysine or acetyl-histidine alone, possibly due to the presence of the charged amino, carboxyl, or carboxylate group adjacent to the acetyl moiety (Figure 3C). Taken together, these results strongly suggest that the P/CAF

bromodomain can interact with acetyl-lysine-containing proteins in a specific manner, and that this interaction is localized to the bromodomain hydrophobic cavity.

- To identify the key residues involved in bromodomain-AcK recognition, the NMR structure of the P/CAF bromodomain in complex with acetyl-histamine was elucidated. As anticipated, the acetylated moiety binds in the bromodomain hydrophobic pocket (Figure 4). The intermolecular interactions are largely hydrophobic in nature, with the methyl group of acetyl-histamine making extensive contacts with the side-chains of Val752, Ala757, and Tyr760, and the methylene groups of acetyl-histamine displaying specific NOEs to Val752, Ala757, Tyr760, Tyr802, and Tyr809. No intermolecular NOEs were observed for the imidazole ring of acetyl-histamine. From the spectral analysis it is clear that the structure of the bromodomain is very similar in both the free and complex forms.
- It is worth noting that the bromodomain-AcK recognition is reminiscent of the interactions between the histone acetyltransferase Hat1 and acetyl-CoA. Although the binding pockets of these two otherwise structurally unrelated proteins are composed of different secondary structural elements, the nature of acetyl-lysine recognition has striking similarities. In particular, Tyr809, Tyr802, Tyr760, and Val752 in the bromodomain appear to be related to Phe220, Phe261, Val254, and Ile217 of Hat1, respectively, in their interactions with the acetyl moiety. This observation may suggest an evolutionary convergent mechanism of acetyl-lysine recognition between bromodomains and histone acetyltransferases.
- To determine the relative contributions of residues within the hydrophobic cavity in bromodomain-AcK binding, site-directed mutagenesis was used to alter residues Tyr809, Tyr802, Tyr760, and Val752 (Table 7).

**Table 7. Structural and Functional Analysis of the P/CAF Bromodomain Mutants**

5	Bromodomain Proteins	Structural Integrity <sup>a</sup>	H4 AcK-Peptide Binding $K_D$ ( $\mu$ M) <sup>b</sup>
	Wild-Type	++++	346 $\pm$ 54
10	Tyr809Ala	++++	No Binding <sup>c</sup>
	Tyr802Ala	+++	> 10,000 <sup>d</sup>
	Tyr760Ala	+++	> 10,000
15	Val752Ala	++	> 10,000

20 a. The effects of mutations on the structural integrity of the bromodomain were assessed by using the  $^{15}\text{N}$ -HSQC spectra. The amide  $^1\text{H}/^{15}\text{N}$  resonances of the mutant proteins were compared to those of the wild-type bromodomain to determine if the particular mutations lead to global or local structure disruption. Severe line-broadening of the amide resonances would indicate protein conformational exchange due to a decrease of structure stability resulting from point mutations. Structural integrity of the mutant proteins is expressed here relative to that of the wild-type, using the signs of “++++” for as stable as the wild-type, “+++” for mildly destabilized, “++” for moderately destabilized, and “-” for completely unfolded.

30 b. The ligand binding affinity ( $K_D$ ) of the bromodomain proteins was estimated by following chemical shift changes of amide peaks in the  $^{15}\text{N}$ -HSQC spectra as a function of the ligand concentration.

c. No detectable ligand binding observed in the NMR titration.

35 d. Ligand binding affinity was significantly reduced and beyond the limit for reliable measurements by NMR titration.

Substitution of Ala for Tyr809 completely abrogated the bromodomain binding to the lysine-acetylated H4 peptide, while the Tyr802Ala, Tyr760Ala, and Val752Ala mutants had significantly reduced ligand binding affinity. To assess whether these mutations disrupted the overall bromodomain fold, the  $^{15}\text{N}$ -HSQC spectra of the mutants was compared to that of the wild-type protein. For the Tyr809Ala mutant, the amide chemical shifts were only affected for a few residues near the mutation site. However, mutations of the other residues in the hydrophobic binding pocket perturbed the local protein conformation to greater extents, particularly the ZA loop (Table 7). Thus, the NMR structural analysis and the mutagenesis studies show that Tyr809, which is structurally supported by Trp746 and Asn803 (Figure 4), is essential for the bromodomain interaction with the acetyl group of acetyl-lysine, while residues of Tyr802, Tyr760, and Val752 likely play both structural and functional roles in the recognition. These residues are highly conserved throughout the bromodomain family (Figure 1), suggesting that recognition of acetyl-lysine may be a feature of bromodomains, in general. Therefore, Val752, Ala757, Tyr760, Tyr802, Asn803, and Tyr809 are key amino acid residues for the P/CAF bromodomain binding to acetyl-lysine.

**Table 8: Amino Acid Sequences of Bromodomains Identified in Figure 1**

	PROTEIN BD	SEQ ID NO:	GenBank Acc. No.	PROTEIN BD	SEQ ID NO:	GenBank Acc. No.
	hsp/CAF	7	U57317	dmFSH-2	25	
5	hsGCN5	8	U57136	scBDF1-2	26	
	ttP55	9	U47321	hsBR140	27	JC2069
	scGCN5	10	Q03330	hsSMAP	28	X87613
	hsP300	11	A54277	ggPB1-1	29	X90849
	hsCBP	12	S39162	ggPB1-2	30	
10	mmCBP	13	S39161	ggPB1-3	31	
	ceYNJ1	14	P34545	ggPB1-4	32	
	hsCCG1-1	15	P21675	ggPB1-5	33	
	msCCG1-1	16	D26114	spBRO-1	34	S54260
	hsCCG1-2	17		spBRO-2	35	
15	msCCG1-2	18		hsSNF2a	36	S45251
	hsRing3-1	19	P25440	hsBRG1	37	S39039
	hsORFX-1	20	D26362	ggBRM	38	X91638
	dmFSH-1	21	P13709	ggBRG1	39	X91637
	scBDF1-1	22	P35817	hsTIF1b	40	X97548
20	hsRing3-2	23		mmTIF1b	41	X99644
	hsORFX-2	24		mmTIF1a	42	S78219

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and the accompanying figures. Such modifications are intended to fall within the scope of the appended claims.



It is further to be understood that all base sizes or amino acid sizes, and all molecular weight or molecular mass values, given for nucleic acids or polypeptides are approximate, and are provided for description.

- 5 Various publications are cited herein, the disclosures of which are hereby incorporated by reference herein in their entireties.

Table 1

**NMR Chemical  
Shift Assignment  
of the P/CAF  
Bromodomain**

RES\_ID 715  
RES\_TYPE GLY  
SPIN\_SYSTEM\_ID 1  
HETEROGENEITY 100  
END\_RES\_DEF

RES\_ID 716  
RES\_TYPE SER  
SPIN\_SYSTEM\_ID 2  
HETEROGENEITY 100  
END\_RES\_DEF

RES\_ID 717  
RES\_TYPE HIS  
SPIN\_SYSTEM\_ID 3  
HETEROGENEITY 100  
END\_RES\_DEF

RES\_ID 718  
RES\_TYPE MET  
SPIN\_SYSTEM\_ID 4  
HETEROGENEITY 100  
END\_RES\_DEF

RES\_ID 719  
RES\_TYPE SER  
SPIN\_SYSTEM\_ID 5  
HETEROGENEITY 100  
END\_RES\_DEF

RES\_ID 720  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 6  
HETEROGENEITY 100

CA 56.296000  
HA 4.361000  
CB 33.140000  
HB1 1.882000  
HB2 1.684000  
CG 25.430000  
HG1 1.585000  
HG2 1.433000  
CD 29.834000  
HD1 1.703000  
CE 41.960000  
HE1 3.003000  
END\_RES\_DEF

RES\_ID 721  
RES\_TYPE GLU  
SPIN\_SYSTEM\_ID 7  
HETEROGENEITY 100

N 122.990000  
HN 8.317000  
CA 54.620000  
HA 4.540000  
CB 29.830000  
HB1 2.024000  
HB2 1.893000  
CG 35.893000  
HG1 2.271000  
END\_RES\_DEF

RES\_ID 722  
RES\_TYPE PRO  
SPIN\_SYSTEM\_ID 8  
HETEROGENEITY 100

CA 63.430000  
HA 4.393000  
CB 32.030000  
HB1 2.224000  
HB2 1.880000  
CG 27.630000  
HG1 2.028000  
CD 50.760000  
HD2 3.656000  
HD1 3.800000  
END\_RES\_DEF

RES\_ID 723  
RES\_TYPE ARG  
SPIN\_SYSTEM\_ID 9

HETEROGENEITY 100  
N 121.192000  
HN 8.416000  
CA 63.430000  
HA 4.331000  
CB 30.930000  
HB1 1.815000  
HB2 1.762000  
CG 27.630000  
HG1 1.681000  
CD 43.603000  
HD1 3.161000  
END\_RES\_DEF

RES\_ID 724  
RES\_TYPE ASP  
SPIN\_SYSTEM\_ID 10  
HETEROGENEITY 100

N 122.012000  
HN 8.273000  
CA 52.415000  
HA 4.874000  
CB 41.400000  
HB1 2.754000  
HB2 2.692000  
END\_RES\_DEF

RES\_ID 725  
RES\_TYPE PRO  
SPIN\_SYSTEM\_ID 11  
HETEROGENEITY 100

CA 65.080000  
HA 4.329000  
CB 32.590000  
HB1 2.326000  
HB2 1.973000  
CG 27.632000  
HG1 2.028000  
CD 51.310000  
HD1 3.866000  
END\_RES\_DEF

RES\_ID 726  
RES\_TYPE ASP  
SPIN\_SYSTEM\_ID 12  
HETEROGENEITY 100

N 119.716000  
HN 8.397000  
CA 55.720000  
HA 4.692000  
CB 40.550000  
HB1 2.792000  
HB2 2.730000  
END\_RES\_DEF

RES\_ID 727  
RES\_TYPE GLN  
SPIN\_SYSTEM\_ID 13  
HETEROGENEITY 100

N 121.356000  
HN 8.196000  
CA 55.920000  
HA 4.163000  
CB 28.730000  
HB1 2.148000  
HB2 1.360000  
CG 24.880000  
HG1 1.280000  
HG2 2.371000  
END\_RES\_DEF

RES\_ID 728  
RES\_TYPE LEU  
SPIN\_SYSTEM\_ID 14  
HETEROGENEITY 100

N 121.356000  
HN 8.210000  
CA 58.473000  
HA 4.045000  
CB 41.400000  
HB1 1.847000  
HB2 1.555000  
CG 27.080000  
HG 1.480000  
CD1 25.970000  
HD1# 0.794000  
CD2 23.226000  
HD2# 0.786000  
END\_RES\_DEF

RES\_ID 729  
RES\_TYPE TYR  
SPIN\_SYSTEM\_ID 15  
HETEROGENEITY 100

N 119.060000  
HN 8.021000

CA 62.320000  
HA 4.038000  
CB 38.640000  
HB1 3.211000  
HB2 3.024000  
CD1 134.350000  
HD1 7.053000  
CE1 119.481000  
HE1 6.882000  
END\_RES\_DEF

RES\_ID 730  
RES\_TYPE SER  
SPIN\_SYSTEM\_ID 16  
HETEROGENEITY 100

N 112.173000  
HN 8.167000  
HA 3.920000  
HB1 3.995000  
END\_RES\_DEF

RES\_ID 731  
RES\_TYPE THR  
SPIN\_SYSTEM\_ID 17  
HETEROGENEITY 100

N 120.372000  
HN 8.059000  
CA 66.730000  
HA 3.924000  
CB 68.930000  
HB 4.247000  
CG2 21.570000  
HG2# 1.142000  
END\_RES\_DEF

RES\_ID 732  
RES\_TYPE LEU  
SPIN\_SYSTEM\_ID 18  
HETEROGENEITY 100

N 120.536000  
HN 8.460000  
CA 57.920000  
HA 3.289000  
CB 39.750000  
HB1 1.532000  
HB2 0.294000  
CG 24.880000  
HG 1.683000  
CD1 25.429000  
HD1# 0.469000  
CD2 19.921000  
HD2# -0.193000  
END\_RES\_DEF

RES\_ID 733  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 19  
HETEROGENEITY 100

N 118.568000  
HN 8.563000  
CA 60.125000  
HA 3.679000  
CB 32.588000  
HB1 1.729000  
HB2 1.360000  
CG 24.880000  
HG1 1.280000  
CD 29.835000  
HD1 1.585000  
CB 41.960000  
HE1 2.918000  
END\_RES\_DEF

RES\_ID 734  
RES\_TYPE SER  
SPIN\_SYSTEM\_ID 20  
HETEROGENEITY 100

N 113.157000  
HN 7.540000  
CA 61.227000  
HA 4.281000  
CB 63.879000  
HB1 4.060000  
END\_RES\_DEF

RES\_ID 735  
RES\_TYPE ILE  
SPIN\_SYSTEM\_ID 21  
HETEROGENEITY 100

N 120.700000  
HN 7.951000  
CA 65.080000  
HA 3.786000  
CB 38.095000  
HB 1.879000

CG1 28.733000  
HG11 1.748000  
HG12 1.052000  
CG2 17.168000  
HG2# 1.003000  
CD1 13.863000  
HD1# 0.619000  
END\_RES\_DEF

RES\_ID 736  
RES\_TYPE LEU  
SPIN\_SYSTEM\_ID 22  
HETEROGENEITY 100

N 119.880000  
HN 8.841000  
CA 58.473000  
HA 4.090000  
CB 41.950000  
HB1 2.090000  
HB2 1.703000  
CG 27.330000  
HG 1.759000  
CD1 26.530000  
HD1# 1.061000  
CD2 23.776000  
HD2# 0.977000  
END\_RES\_DEF

RES\_ID 737  
RES\_TYPE GLN  
SPIN\_SYSTEM\_ID 23  
HETEROGENEITY 100

N 117.256000  
HN 8.505000  
CA 59.020000  
HA 4.032000  
CB 28.182000  
HB1 2.327000  
HB2 2.263000  
CG 34.240000  
HG1 2.536000  
HG2 2.461000  
END\_RES\_DEF

RES\_ID 738  
RES\_TYPE GLN  
SPIN\_SYSTEM\_ID 24  
HETEROGENEITY 100

N 118.896000  
HN 8.033000  
CA 59.574000  
HA 4.196000  
CB 29.835000  
HB1 2.482000  
HB2 2.469000  
CG 35.342000  
HG1 2.840000  
HG2 2.467000  
NE2 110.369000  
HE21 7.022000  
HE22 6.916000  
END\_RES\_DEF

RES\_ID 739  
RES\_TYPE VAL  
SPIN\_SYSTEM\_ID 25  
HETEROGENEITY 100

N 119.716000  
HN 8.526000  
CA 67.830000  
HA 3.844000  
CB 32.030000  
HB 2.384000  
CG1 23.330000  
HG1# 1.183000  
CG2 22.120000  
HG2# 1.033000  
END\_RES\_DEF

RES\_ID 740  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 26  
HETEROGENEITY 100

N 114.633000  
HN 8.572000  
CA 59.574000  
HA 3.886000  
CB 32.380000  
HB1 1.873000  
HG1 1.022000  
HD1 1.520000  
END\_RES\_DEF

RES\_ID 741  
RES\_TYPE SER

SPIN\_SYSTEM\_ID 27  
 HETEROGENEITY 100  
 N 110.369000  
 HN 7.557000  
 CA 59.024000  
 HA 4.448000  
 CB 63.980000  
 HB1 4.004000  
 END\_RES\_DEF

RES\_ID 742  
 RES\_TYPE HIS  
 SPIN\_SYSTEM\_ID 28  
 HETEROGENEITY 100

N 125.619000  
 HN 7.536000  
 CA 58.473000  
 HA 3.967000  
 CB 32.588000  
 HB1 2.990000  
 HB2 2.799000  
 CD2 118.930000  
 HD2 4.978000  
 CE1 118.755000  
 HE1 7.522000  
 END\_RES\_DEF

RES\_ID 743  
 RES\_TYPE GLN  
 SPIN\_SYSTEM\_ID 29  
 HETEROGENEITY 100

N 128.571000  
 HN 8.543000  
 CA 59.125000  
 HA 4.209000  
 CB 29.834000  
 HB1 2.111000  
 CG 33.690000  
 HG1 2.390000  
 NE2 112.173000  
 HE21 7.581000  
 HE22 6.870000  
 END\_RES\_DEF

RES\_ID 744  
 RES\_TYPE SER  
 SPIN\_SYSTEM\_ID 30  
 HETEROGENEITY 100

N 119.060000  
 HN 11.668000  
 CA 60.125000  
 HA 4.838000  
 CB 63.980000  
 HB1 4.334000  
 HB2 3.926000  
 END\_RES\_DEF

RES\_ID 745  
 RES\_TYPE ALA  
 SPIN\_SYSTEM\_ID 31  
 HETEROGENEITY 100

N 117.584000  
 HN 7.868000  
 CA 53.510000  
 HA 4.396000  
 CB 20.470000  
 HB# 1.688000  
 END\_RES\_DEF

RES\_ID 746  
 RES\_TYPE TRP  
 SPIN\_SYSTEM\_ID 32  
 HETEROGENEITY 100

N 116.600000  
 HN 7.135000  
 CA 60.691000  
 HA 4.368000  
 CB 27.630000  
 HB1 3.594000  
 HB2 3.351000  
 CD1 128.843000  
 HD1 7.897000  
 NE1 110.861000  
 HE1 10.474000  
 CE3 122.234000  
 HE3 7.336000  
 CZ2 116.177000  
 HZ2 7.382000  
 CZ3 123.336000  
 HZ3 7.197000  
 CH2 126.089000  
 HH2 7.150000  
 END\_RES\_DEF

RES\_ID 747

RES\_TYPE PRO  
 SPIN\_SYSTEM\_ID 33  
 HETEROGENEITY 100  
 CA 64.531000  
 HA 3.756000  
 CB 29.835000  
 HB1 0.487000  
 HB2 -0.783000  
 CG 26.530000  
 HG1 0.233000  
 HG2 -0.931000  
 CD 50.212000  
 HD2 1.567000  
 HD1 2.177000  
 END\_RES\_DEF

RES\_ID 748  
 RES\_TYPE PHE  
 SPIN\_SYSTEM\_ID 34  
 HETEROGENEITY 100  
 N 113.321000  
 HN 7.585000  
 CA 55.719000  
 HA 4.930000  
 CB 39.202000  
 HB1 3.491000  
 HB2 2.532000  
 CD1 133.248000  
 HD1 7.099000  
 HE1 7.174000  
 HZ 7.296000  
 END\_RES\_DEF

RES\_ID 749  
 RES\_TYPE MET  
 SPIN\_SYSTEM\_ID 35  
 HETEROGENEITY 100  
 N 117.748000  
 HN 7.115000  
 CA 56.820000  
 HA 4.286000  
 CB 32.590000  
 HB1 2.233000  
 HB2 2.174000  
 CG 33.140000  
 HG1 2.851000  
 CE 17.168000  
 HE# 2.175000  
 END\_RES\_DEF

RES\_ID 750  
 RES\_TYPE GLU  
 SPIN\_SYSTEM\_ID 36  
 HETEROGENEITY 100  
 N 113.813000  
 HN 7.709000  
 CA 53.516000  
 HA 4.849000  
 CB 31.487000  
 HB1 2.091000  
 HB2 1.730000  
 CG 35.893000  
 HG1 2.164000  
 END\_RES\_DEF

RES\_ID 751  
 RES\_TYPE PRO  
 SPIN\_SYSTEM\_ID 37  
 HETEROGENEITY 100  
 CA 62.879000  
 HA 4.242000  
 CB 32.040000  
 HB1 2.328000  
 HB2 1.683000  
 CG 27.080000  
 HG1 2.126000  
 HG2 1.978000  
 CD 50.763000  
 HD1 3.670000  
 END\_RES\_DEF

RES\_ID 752  
 RES\_TYPE VAL  
 SPIN\_SYSTEM\_ID 38  
 HETEROGENEITY 100  
 N 124.450000  
 HN 8.124000  
 CA 63.430000  
 HA 3.553000  
 CB 32.580000  
 HB 1.145000  
 CG1 21.573000  
 HG1# 0.464000  
 CG2 21.573000  
 HG2# 0.169000

END\_RES\_DEF  
 RES\_ID 753  
 RES\_TYPE LYS  
 SPIN\_SYSTEM\_ID 39  
 HETEROGENEITY 100  
 N 129.883000  
 HN 9.045000  
 CA 56.310000  
 HA 4.370000  
 CB 32.880000  
 HB1 1.873000  
 HG1 1.435000  
 HD1 1.673000  
 HE1 2.985000  
 END\_RES\_DEF

RES\_ID 754  
 RES\_TYPE ARG  
 SPIN\_SYSTEM\_ID 40  
 HETEROGENEITY 100  
 N 120.208000  
 HN 8.054000  
 END\_RES\_DEF  
 RES\_ID 755  
 RES\_TYPE THR  
 SPIN\_SYSTEM\_ID 41  
 HETEROGENEITY 100  
 CA 63.430000  
 HA 4.038000  
 CB 68.380000  
 HB 4.293000  
 CG2 22.670000  
 HG2# 1.267000  
 END\_RES\_DEF

RES\_ID 756  
 RES\_TYPE GLU  
 SPIN\_SYSTEM\_ID 42  
 HETEROGENEITY 100  
 N 118.732000  
 HN 7.209000  
 CA 56.270000  
 HA 4.448000  
 CB 30.930000  
 HB1 2.174000  
 HB2 2.000000  
 CG 36.440000  
 HG1 2.292000  
 END\_RES\_DEF

RES\_ID 757  
 RES\_TYPE ALA  
 SPIN\_SYSTEM\_ID 43  
 HETEROGENEITY 100  
 N 122.504000  
 HN 7.379000  
 CA 50.220000  
 HA 4.937000  
 CB 19.370000  
 HB# 1.082000  
 END\_RES\_DEF

RES\_ID 758  
 RES\_TYPE PRO  
 SPIN\_SYSTEM\_ID 44  
 HETEROGENEITY 100  
 CA 65.080000  
 HA 4.496000  
 CB 31.487000  
 HB1 2.374000  
 HB2 2.027000  
 CG 27.632000  
 HG1 2.122000  
 HG2 2.038000  
 CD 50.212000  
 HD2 3.515000  
 HD1 3.717000  
 END\_RES\_DEF

RES\_ID 759  
 RES\_TYPE GLY  
 SPIN\_SYSTEM\_ID 45  
 HETEROGENEITY 100  
 END\_RES\_DEF  
 RES\_ID 760  
 RES\_TYPE TYR  
 SPIN\_SYSTEM\_ID 46  
 HETEROGENEITY 100  
 N 122.504000  
 HN 7.945000  
 CA 62.328000  
 HA 3.536000

CB 39.750000  
 HB1 2.689000  
 HB2 2.487000  
 CD1 133.799000  
 HD1 5.120000  
 CE1 118.379000  
 HE1 6.070000  
 END\_RES\_DEF

RES\_ID 761  
 RES\_TYPE TYR  
 SPIN\_SYSTEM\_ID 47  
 HETEROGENEITY 100  
 N 113.157000  
 HN 8.225000  
 CA 60.676000  
 HA 4.101000  
 CB 37.550000  
 HB1 3.189000  
 HB2 2.801000  
 CD1 134.901000  
 HD1 7.342000  
 CE1 118.930000  
 HE1 6.646000  
 END\_RES\_DEF

RES\_ID 762  
 RES\_TYPE GLU  
 SPIN\_SYSTEM\_ID 48  
 HETEROGENEITY 100  
 N 117.912000  
 HN 7.702000  
 CA 57.922000  
 HA 4.209000  
 CB 29.480000  
 HB1 2.086000  
 CG 37.545000  
 HG1 2.325000  
 HG2 2.265000  
 END\_RES\_DEF

RES\_ID 763  
 RES\_TYPE VAL  
 SPIN\_SYSTEM\_ID 49  
 HETEROGENEITY 100  
 N 115.453000  
 HN 7.135000  
 CA 63.430000  
 HA 4.077000  
 CB 33.690000  
 HB 2.015000  
 CG1 21.020000  
 HG1# 1.045000  
 CG2 21.574000  
 HG2# 0.991000  
 END\_RES\_DEF

RES\_ID 764  
 RES\_TYPE ILE  
 SPIN\_SYSTEM\_ID 50  
 HETEROGENEITY 100  
 N 122.832000  
 HN 7.947000  
 CA 57.920000  
 HA 3.916000  
 CB 34.240000  
 HB 1.205000  
 CG1 24.878000  
 HG11 0.798000  
 HG12 0.216000  
 CG2 16.617000  
 HG2# 0.380000  
 CD1 9.457000  
 HD1# 0.537000  
 END\_RES\_DEF

RES\_ID 765  
 RES\_TYPE ARG  
 SPIN\_SYSTEM\_ID 51  
 HETEROGENEITY 100  
 N 125.291000  
 HN 7.749000  
 CA 57.371000  
 HA 3.875000  
 CB 30.936000  
 HB1 1.388000  
 HB2 1.211000  
 CG 27.080000  
 HG1 1.319000  
 HG2 1.173000  
 CD 43.052000  
 HD1 2.971000  
 END\_RES\_DEF

RES\_ID 766

RES\_TYPE SER  
SPIN\_SYSTEM\_ID 52  
HETEROGENEITY 100  
N 116.600000  
HN 8.387000  
CA 54.618000  
HA 4.984000  
CB 38.640000  
HB1 3.034000  
HB2 2.907000  
END\_RES\_DEF

RES\_ID 767  
RES\_TYPE PRO  
SPIN\_SYSTEM\_ID 53  
HETEROGENEITY 100  
CA 63.429000  
HA 4.083000  
CB 32.588000  
HB1 2.209000  
CG 28.180000  
HG1 2.177000  
HG2 1.883000  
CD 50.763000  
HD2 3.390000  
HD1 3.623000  
END\_RES\_DEF

RES\_ID 768  
RES\_TYPE MET  
SPIN\_SYSTEM\_ID 54  
HETEROGENEITY 100  
N 119.060000  
HN 8.430000  
CA 54.067000  
HA 4.935000  
CB 31.487000  
HB1 1.989000  
HB2 1.353000  
CG 30.930000  
HG1 2.690000  
CE 14.414000  
HE# 1.929000  
END\_RES\_DEF

RES\_ID 769  
RES\_TYPE ASP  
SPIN\_SYSTEM\_ID 55  
HETEROGENEITY 100  
N 119.060000  
HN 7.365000  
CA 53.516000  
HA 4.745000  
CB 44.154000  
HB1 2.371000  
END\_RES\_DEF

RES\_ID 770  
RES\_TYPE LEU  
SPIN\_SYSTEM\_ID 56  
HETEROGENEITY 100  
N 116.272000  
HN 9.055000  
CA 57.922000  
HA 4.036000  
CB 41.400000  
HB1 2.095000  
HB2 1.395000  
CG 27.080000  
HG 1.713000  
CD1 27.080000  
HD1# 0.940000  
CD2 22.675000  
HD2# 0.628000  
END\_RES\_DEF

RES\_ID 771  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 57  
HETEROGENEITY 100  
N 128.079000  
HN 8.738000  
CA 60.676000  
HA 4.198000  
CB 32.037000  
HB1 2.330000  
HB2 2.224000  
CG 25.280000  
HG1 1.483000  
HG2 1.403000  
CD 30.385000  
HD1 1.793000  
HD2 1.696000  
CE 41.950000  
HE1 2.965000

END\_RES\_DEF  
RES\_ID 772  
RES\_TYPE THR  
SPIN\_SYSTEM\_ID 58  
HETEROGENEITY 100  
N 122.176000  
HN 9.445000  
CA 67.040000  
HA 3.845000  
CB 67.835000  
HB 4.090000  
CG2 22.124000  
HG2# 1.058000  
END\_RES\_DEF

RES\_ID 773  
RES\_TYPE MET  
SPIN\_SYSTEM\_ID 59  
HETEROGENEITY 100  
N 117.912000  
HN 7.882000  
CA 60.676000  
HA 4.319000  
CB 33.342000  
HB1 2.093000  
HB2 1.915000  
CG 33.139000  
HG1 2.621000  
HG2 2.496000  
CE 16.620000  
HE# 1.241000  
END\_RES\_DEF

RES\_ID 774  
RES\_TYPE SER  
SPIN\_SYSTEM\_ID 60  
HETEROGENEITY 100  
N 116.108000  
HN 7.958000  
CA 62.879000  
HA 4.200000  
CB 62.879000  
HB1 4.368000  
HB2 4.040000  
END\_RES\_DEF

RES\_ID 775  
RES\_TYPE GLU  
SPIN\_SYSTEM\_ID 61  
HETEROGENEITY 100  
N 124.471000  
HN 8.150000  
CA 59.570000  
HA 4.045000  
CB 29.280000  
HB1 2.246000  
HB2 2.063000  
CG 36.443000  
HG1 2.345000  
HG2 2.176000  
END\_RES\_DEF

RES\_ID 776  
RES\_TYPE ARG  
SPIN\_SYSTEM\_ID 62  
HETEROGENEITY 100  
N 120.372000  
HN 8.391000  
CA 60.676000  
HA 3.869000  
CB 30.385000  
HB1 2.047000  
HB2 1.076000  
CG 29.284000  
HG1 1.722000  
HG2 0.877000  
CD 44.154000  
HD1 2.578000  
HD2 2.051000  
END\_RES\_DEF

RES\_ID 777  
RES\_TYPE LEU  
SPIN\_SYSTEM\_ID 63  
HETEROGENEITY 100  
N 120.208000  
HN 8.856000  
CA 58.470000  
HA 4.691000  
CB 42.621000  
HB1 2.295000  
HB2 1.925000  
CG 27.080000  
HG 1.832000

CD1 25.429000  
HD1# 1.067000  
CD2 27.081000  
HD2# 0.871000  
END\_RES\_DEF

RES\_ID 778  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 64  
HETEROGENEITY 100  
N 120.372000  
HN 7.958000  
CA 59.574000  
HA 4.333000  
CB 32.588000  
HB1 2.055000  
CG 24.878000  
HG1 1.596000  
CD 29.835000  
HD1 1.804000  
CE 41.951000  
HE1 2.990000  
END\_RES\_DEF

RES\_ID 779  
RES\_TYPE ASN  
SPIN\_SYSTEM\_ID 65  
HETEROGENEITY 100  
N 116.108000  
HN 7.947000  
CA 53.510000  
HA 4.771000  
CB 38.095000  
HB1 3.019000  
HB2 2.773000  
ND2 112.665000  
HD21 7.598000  
HD22 6.969000  
END\_RES\_DEF

RES\_ID 780  
RES\_TYPE ARG  
SPIN\_SYSTEM\_ID 66  
HETEROGENEITY 100  
N 114.141000  
HN 8.158000  
CA 56.821000  
HA 4.405000  
CB 25.429000  
HB1 2.097000  
HB2 2.022000  
CG 27.632000  
HG1 1.539000  
HG2 1.534000  
CD 43.050000  
HD1 3.060000  
HD2 3.024000  
END\_RES\_DEF

RES\_ID 781  
RES\_TYPE TYR  
SPIN\_SYSTEM\_ID 67  
HETEROGENEITY 100  
N 116.764000  
HN 8.222000  
CA 60.125000  
HA 4.064000  
CB 40.850000  
HB1 2.948000  
HB2 2.055000  
CD1 134.350000  
HD1 6.285000  
CE1 118.930000  
HE1 6.709000  
END\_RES\_DEF

RES\_ID 782  
RES\_TYPE TYR  
SPIN\_SYSTEM\_ID 68  
HETEROGENEITY 100  
N 114.633000  
HN 8.014000  
CA 57.920000  
HA 4.528000  
CB 36.443000  
HB1 3.062000  
HB2 2.907000  
CD1 133.248000  
HD1 7.175000  
CE1 120.582000  
HE1 7.286000  
END\_RES\_DEF

RES\_ID 783  
RES\_TYPE VAL

SPIN\_SYSTEM\_ID 69  
HETEROGENEITY 100  
N 115.780000  
HN 7.698000  
CA 62.330000  
HA 4.083000  
CB 31.500000  
HB 2.321000  
CG1 21.570000  
HG1# 0.944000  
CG2 18.820000  
HG2# 0.823000  
END\_RES\_DEF

RES\_ID 784  
RES\_TYPE SER  
SPIN\_SYSTEM\_ID 70  
HETEROGENEITY 100  
N 111.353000  
HN 7.415000  
CA 55.719000  
HA 4.741000  
CB 66.183000  
HB1 4.200000  
HB2 3.750000  
END\_RES\_DEF

RES\_ID 785  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 71  
HETEROGENEITY 100  
CA 59.030000  
HA 4.021000  
CB 31.590000  
END\_RES\_DEF

RES\_ID 786  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 72  
HETEROGENEITY 100  
N 120.208000  
HN 8.244000  
CA 59.720000  
HA 4.062000  
CB 30.385000  
HB1 1.779000  
CG 24.530000  
CD 28.182000  
HD1 1.680000  
CE 41.670000  
HE1 3.137000  
HE2 3.045000  
END\_RES\_DEF

RES\_ID 787  
RES\_TYPE LEU  
SPIN\_SYSTEM\_ID 73  
HETEROGENEITY 100  
N 118.732000  
HN 7.422000  
CA 57.922000  
HA 4.213000  
CB 43.603000  
HB1 1.996000  
HB2 1.891000  
CG 27.632000  
HG 1.794000  
CD1 25.979000  
HD1# 0.924000  
CD2 23.776000  
HD2# 0.895000  
END\_RES\_DEF

RES\_ID 788  
RES\_TYPE PHE  
SPIN\_SYSTEM\_ID 74  
HETEROGENEITY 100  
N 118.732000  
HN 6.928000  
CA 60.676000  
HA 3.763000  
CB 39.750000  
HB1 2.945000  
HB2 2.381000  
CD1 133.799000  
HD1 6.400000  
CE1 131.596000  
HE1 6.928000  
END\_RES\_DEF

RES\_ID 789  
RES\_TYPE MET  
SPIN\_SYSTEM\_ID 75  
HETEROGENEITY 100  
N 116.272000

HN 8.489000  
CA 59.020000  
HA 3.911000  
CB 32.590000  
HB1 2.318000  
HB2 2.208000  
CG 33.140000  
HG1 2.942000  
HG2 2.611000  
CE 17.168000  
HE# 2.027000  
END\_RES\_DEF

RES\_ID 790  
RES\_TYPE ALA  
SPIN\_SYSTEM\_ID 76  
HETEROGENEITY 100  
N 119.716000  
HN 8.000000  
CA 55.170000  
HA 4.084000  
CB 18.270000  
HB# 1.485000  
END\_RES\_DEF

RES\_ID 791  
RES\_TYPE ASP  
SPIN\_SYSTEM\_ID 77  
HETEROGENEITY 100  
N 119.716000  
HN 7.376000  
CA 57.371000  
HA 4.371000  
CB 38.646000  
HB1 2.730000  
END\_RES\_DEF

RES\_ID 792  
RES\_TYPE LEU  
SPIN\_SYSTEM\_ID 78  
HETEROGENEITY 100  
N 119.550000  
HN 7.363000  
CA 57.922000  
HA 3.398000  
CB 40.299000  
HB1 0.757000  
HB2 0.442000  
CG 27.632000  
HG 0.707000  
CD1 24.327000  
HD1# 0.184000  
CD2 25.979000  
HD2# 0.061000  
END\_RES\_DEF

RES\_ID 793  
RES\_TYPE GLN  
SPIN\_SYSTEM\_ID 79  
HETEROGENEITY 100  
N 114.141000  
HN 8.069000  
CA 59.024000  
HA 3.804000  
CB 28.733000  
HB1 2.157000  
HB2 2.097000  
CG 35.342000  
HG1 2.460000  
NE2 111.353000  
HE21 7.319000  
HE22 7.222000  
END\_RES\_DEF

RES\_ID 794  
RES\_TYPE ARG  
SPIN\_SYSTEM\_ID 80  
HETEROGENEITY 100  
N 118.568000  
HN 7.382000  
CA 58.473000  
HA 4.078000  
CB 29.835000  
HB1 1.973000  
HB2 1.886000  
CG 27.080000  
HG1 1.742000  
CD 43.603000  
HD1 3.390000  
HD2 3.325000  
END\_RES\_DEF

RES\_ID 795  
RES\_TYPE VAL  
SPIN\_SYSTEM\_ID 81

HETEROGENEITY 100  
N 117.912000  
HN 7.013000  
CA 66.730000  
HA 3.039000  
CB 30.930000  
HB 1.435000  
CG1 22.124000  
HG1# 0.479000  
CG2 21.573000  
HG2# 0.142000  
END\_RES\_DEF

RES\_ID 796  
RES\_TYPE PHE  
SPIN\_SYSTEM\_ID 82  
HETEROGENEITY 100  
N 116.928000  
HN 6.357000  
CA 58.470000  
HA 4.161000  
CB 38.096000  
HB1 3.090000  
HB2 2.944000  
CD1 132.147000  
HD1 6.641000  
CE1 131.596000  
HE1 6.456000  
CZ 129.393000  
HZ 6.406000  
END\_RES\_DEF

RES\_ID 797  
RES\_TYPE THR  
SPIN\_SYSTEM\_ID 83  
HETEROGENEITY 100  
N 115.289000  
HN 9.047000  
CA 66.734000  
HA 3.838000  
CB 68.380000  
HB 4.210000  
CG2 22.120000  
HG2# 1.296000  
END\_RES\_DEF

RES\_ID 798  
RES\_TYPE ASN  
SPIN\_SYSTEM\_ID 84  
HETEROGENEITY 100  
N 120.700000  
HN 8.846000  
CA 55.170000  
HA 4.315000  
CB 38.090000  
HB1 2.985000  
HB2 2.661000  
END\_RES\_DEF

RES\_ID 799  
RES\_TYPE CYS  
SPIN\_SYSTEM\_ID 85  
HETEROGENEITY 100  
N 116.928000  
HN 6.893000  
CA 62.157000  
HA 4.405000  
CB 26.530000  
HB1 3.304000  
HB2 3.032000  
END\_RES\_DEF

RES\_ID 800  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 86  
HETEROGENEITY 100  
N 116.764000  
HN 7.799000  
CA 58.473000  
HA 4.204000  
CB 32.588000  
HB1 1.743000  
CG 25.429000  
HG1 1.313000  
HG2 0.138000  
CD 29.835000  
HD1 1.291000  
CE 41.400000  
HE1 2.486000  
HE2 2.421000  
END\_RES\_DEF

RES\_ID 801  
RES\_TYPE GLU  
SPIN\_SYSTEM\_ID 87

HETEROGENEITY 100  
N 117.912000  
HN 7.945000  
CA 57.992000  
HA 4.250000  
CB 30.385000  
HB1 2.172000  
HB2 2.003000  
CG 36.994000  
HG1 2.407000  
HG2 2.203000  
END\_RES\_DEF

RES\_ID 802  
RES\_TYPE TYR  
SPIN\_SYSTEM\_ID 88  
HETEROGENEITY 100  
N 116.600000  
HN 7.744000  
CA 60.676000  
HA 4.369000  
CB 41.400000  
HB1 2.929000  
CD1 134.901000  
HD1 6.989000  
CE1 119.481000  
HE1 6.823000  
END\_RES\_DEF

RES\_ID 803  
RES\_TYPE ASN  
SPIN\_SYSTEM\_ID 89  
HETEROGENEITY 100  
N 115.944000  
HN 8.241000  
CA 51.864000  
HA 5.024000  
CB 40.849000  
HB1 3.069000  
HB2 2.907000  
ND2 118.732000  
HD21 8.316000  
HD22 7.809000  
END\_RES\_DEF

RES\_ID 804  
RES\_TYPE ALA  
SPIN\_SYSTEM\_ID 90  
HETEROGENEITY 100  
END\_RES\_DEF  
RES\_ID 805  
RES\_TYPE PRO  
SPIN\_SYSTEM\_ID 91  
HETEROGENEITY 100  
CA 63.980000  
HA 2.422000  
HB1 1.949000  
HG1 1.648000  
HG2 1.558000  
CD 50.762000  
HD2 3.601000  
HD1 3.706000  
END\_RES\_DEF

RES\_ID 806  
RES\_TYPE GLU  
SPIN\_SYSTEM\_ID 92  
HETEROGENEITY 100  
N 112.993000  
HN 8.246000  
CA 56.820000  
HA 4.185000  
CB 28.733000  
HB1 2.095000  
HB2 1.973000  
CG 36.270000  
HG1 2.200000  
END\_RES\_DEF

RES\_ID 807  
RES\_TYPE SER  
SPIN\_SYSTEM\_ID 93  
HETEROGENEITY 100  
N 115.780000  
HN 8.112000  
CA 58.473000  
HA 4.406000  
CB 66.183000  
HB1 4.393000  
HB2 4.157000  
END\_RES\_DEF

RES\_ID 808  
RES\_TYPE GLU

SPIN\_SYSTEM\_ID 94  
HETEROGENEITY 100  
N 123.488000  
HN 9.061000  
CA 59.574000  
HA 4.232000  
CB 29.835000  
HB1 2.169000  
CG 36.443000  
HG1 2.528000  
END\_RES\_DEF

RES\_ID 809  
RES\_TYPE TYR  
SPIN\_SYSTEM\_ID 95  
HETEROGENEITY 100  
N 116.436000  
HN 8.072000  
CA 60.120000  
HA 3.834000  
CB 37.550000  
HB1 3.018000  
HB2 2.738000  
CD1 132.698000  
HD1 6.891000  
CE1 120.032000  
HE1 7.011000  
END\_RES\_DEF

RES\_ID 810  
RES\_TYPE TYR  
SPIN\_SYSTEM\_ID 96  
HETEROGENEITY 100  
N 119.880000  
HN 7.356000  
CA 61.777000  
HA 3.819000  
CB 40.300000  
HB1 3.390000  
HB2 2.500000  
CD1 136.553000  
HD1 7.094000  
CE1 119.481000  
HE1 7.000000  
END\_RES\_DEF

RES\_ID 811  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 97  
HETEROGENEITY 100  
N 118.076000  
HN 8.072000  
CA 60.676000  
HA 4.204000  
CB 32.588000  
HB1 2.091000  
CG 25.979000  
HG1 1.819000  
HG2 1.582000  
CD 29.834000  
HD1 1.813000  
CE 41.963000  
HE1 2.962000  
END\_RES\_DEF

RES\_ID 812  
RES\_TYPE CYS  
SPIN\_SYSTEM\_ID 98  
HETEROGENEITY 100  
N 116.764000  
HN 8.520000  
CA 65.087000  
HA 4.202000  
CB 27.080000  
HB1 3.396000  
HB2 3.056000  
END\_RES\_DEF

RES\_ID 813  
RES\_TYPE ALA  
SPIN\_SYSTEM\_ID 99  
HETEROGENEITY 100  
N 120.700000  
HN 8.315000  
CA 55.563000  
HA 3.834000  
CB 18.270000  
HB# 1.597000  
END\_RES\_DEF

RES\_ID 814  
RES\_TYPE ASN  
SPIN\_SYSTEM\_ID 100  
HETEROGENEITY 100  
N 115.453000

HN 8.068000  
CA 56.270000  
HA 4.329000  
CB 38.646000  
HB1 2.877000  
HB2 2.834000  
END\_RES\_DEF

RES\_ID 815  
RES\_TYPE ILE  
SPIN\_SYSTEM\_ID 101  
HETEROGENEITY 100

N 119.880000  
HN 7.912000  
CA 65.080000  
HA 3.646000  
CB 39.197000  
HB 1.924000  
CG1 29.284000  
HG11 1.882000  
HG12 1.201000  
CG2 17.718000  
HG2# 1.017000  
CD1 13.863000  
HD1# 0.940000  
END\_RES\_DEF

RES\_ID 816  
RES\_TYPE LEU  
SPIN\_SYSTEM\_ID 102  
HETEROGENEITY 100

N 122.504000  
HN 8.556000  
CA 56.820000  
HA 3.670000  
CB 41.951000  
HB1 1.405000  
HB2 1.199000  
CG 26.530000  
HG 1.580000  
CD1 24.327000  
HD1# 0.701000  
CD2 25.429000  
HD2# 0.696000  
END\_RES\_DEF

RES\_ID 817  
RES\_TYPE GLU  
SPIN\_SYSTEM\_ID 103  
HETEROGENEITY 100

N 120.700000  
HN 8.073000  
CA 60.125000  
HA 3.185000  
CB 29.835000  
HB1 1.720000  
HB2 1.310000  
CG 37.545000  
HG1 2.001000  
HG2 1.922000  
END\_RES\_DEF

RES\_ID 818  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 104  
HETEROGENEITY 100

N 117.584000  
HN 7.145000  
CA 59.688000  
HA 4.075000  
CB 32.588000  
HB1 1.929000  
CG 25.644000  
HG1 1.492000  
CD 29.284000  
HD1 1.681000  
CE 41.963000  
HE1 2.964000  
END\_RES\_DEF

RES\_ID 819  
RES\_TYPE PHE  
SPIN\_SYSTEM\_ID 105  
HETEROGENEITY 100

N 121.028000  
HN 7.869000  
CA 61.230000  
HA 4.328000  
CB 39.200000  
HB1 3.133000  
HB2 3.047000  
CD1 133.800000  
HD1 7.180000  
END\_RES\_DEF

RES\_ID 820  
RES\_TYPE PHE  
SPIN\_SYSTEM\_ID 106  
HETEROGENEITY 100

N 120.700000  
HN 9.126000  
CA 60.691000  
HA 3.961000  
CB 38.640000  
HB1 3.289000  
HB2 3.067000  
CD1 133.248000  
HD1 6.904000  
CE1 132.698000  
HE1 7.011000  
END\_RES\_DEF

RES\_ID 821  
RES\_TYPE PHE  
SPIN\_SYSTEM\_ID 107  
HETEROGENEITY 100

N 118.076000  
HN 8.359000  
CA 61.770000  
HA 3.840000  
CB 38.090000  
HB1 3.064000  
CD1 133.248000  
HD1 7.175000  
CE1 132.698000  
HE1 7.294000  
CZ 131.596000  
HZ 7.430000  
END\_RES\_DEF

RES\_ID 822  
RES\_TYPE SER  
SPIN\_SYSTEM\_ID 108  
HETEROGENEITY 100

N 114.961000  
HN 7.906000  
CA 61.773000  
HA 4.200000  
CB 62.879000  
HB1 4.007000  
END\_RES\_DEF

RES\_ID 823  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 109  
HETEROGENEITY 100

N 120.864000  
HN 7.938000  
CA 56.820000  
HA 4.008000  
CB 31.487000  
HB1 1.730000  
HB2 1.567000  
CG 23.226000  
HG1 0.833000  
CD 27.080000  
HD1 1.403000  
CE 42.501000  
HE1 2.569000  
HE2 2.422000  
END\_RES\_DEF

RES\_ID 824  
RES\_TYPE ILE  
SPIN\_SYSTEM\_ID 110  
HETEROGENEITY 100

N 116.928000  
HN 8.101000  
CA 64.530000  
HA 3.818000  
CB 36.990000  
HB 1.746000  
CG1 26.530000  
HG11 1.140000  
HG12 1.073000  
CG2 18.820000  
HG2# 0.654000  
CD1 13.312000  
HD1# 0.541000  
END\_RES\_DEF

RES\_ID 825  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 111  
HETEROGENEITY 100

N 122.176000  
HN 7.546000  
CA 59.024000  
HA 4.043000  
CB 32.360000

HB1 1.879000  
HB2 1.757000  
CG 24.878000  
HG1 1.390000  
HG2 1.302000  
CD 29.284000  
HD1 1.633000  
CE 41.400000  
HE1 2.913000  
END\_RES\_DEF

RES\_ID 826  
RES\_TYPE GLU  
SPIN\_SYSTEM\_ID 112  
HETEROGENEITY 100

N 121.192000  
HN 8.063000  
CA 59.024000  
HA 3.995000  
CB 29.834000  
HB1 2.058000  
CG 36.050000  
HG1 2.342000  
HG2 2.205000  
END\_RES\_DEF

RES\_ID 827  
RES\_TYPE ALA  
SPIN\_SYSTEM\_ID 113  
HETEROGENEITY 100

N 117.748000  
HN 7.620000  
CA 52.410000  
HA 4.291000  
CB 19.920000  
HB# 1.358000  
END\_RES\_DEF

RES\_ID 828  
RES\_TYPE GLY  
SPIN\_SYSTEM\_ID 114  
HETEROGENEITY 100

N 126.767000  
HN 7.744000  
CA 45.902000  
HA1 4.019000  
HA2 3.935000  
END\_RES\_DEF

RES\_ID 829  
RES\_TYPE LEU  
SPIN\_SYSTEM\_ID 115  
HETEROGENEITY 100

N 117.912000  
HN 7.742000  
CA 55.719000  
HA 4.215000  
CB 43.052000  
HB1 1.562000  
CG 27.632000  
HG 1.536000  
CD1 23.776000  
HD1# 0.711000  
END\_RES\_DEF

RES\_ID 830  
RES\_TYPE ILE  
SPIN\_SYSTEM\_ID 116  
HETEROGENEITY 100

N 115.453000  
HN 7.458000  
CA 60.676000  
HA 4.232000  
CB 39.748000  
HB 1.810000  
CG1 27.080000  
HG11 1.314000  
HG12 0.918000  
CG2 17.718000  
HG2# 0.815000  
CD1 13.312000  
HD1# 0.794000  
END\_RES\_DEF

RES\_ID 831  
RES\_TYPE ASP  
SPIN\_SYSTEM\_ID 117  
HETEROGENEITY 100

N 123.488000  
HN 8.270000  
CA 54.620000  
HA 4.571000  
CB 41.400000  
HB1 2.693000  
HB2 2.540000

END\_RES\_DEF

RES\_ID 832  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 118  
HETEROGENEITY 100

N 125.450000  
HN 7.774000  
CA 57.720000  
HA 4.082000  
CB 33.410000  
END\_RES\_DEF

Unambiguous NOE-derived Inter-proton Distance Restraints

[illegible]





3

(( segid "BRD" \* and resid 24 and name HA ))  
 ASST [ 1671 ] 2.800 2.000 2.000 peak 1661 weight 0.11000E+01 volume 0.41492E+03 ppm1 8.661 ppm2 4.775  
 (( segid "BRD" \* and resid 24 and name HN ))  
 (( segid "BRD" \* and resid 24 and name H1 ))  
 ASST [ 1672 ] 2.800 2.000 2.000 peak 1671 weight 0.11000E+01 volume 0.42165E+03 ppm1 8.660 ppm2 3.454  
 (( segid "BRD" \* and resid 24 and name H2 ))  
 (( segid "BRD" \* and resid 24 and name H3 ))  
 ASST [ 1673 ] 2.800 2.000 2.000 peak 1681 weight 0.11000E+01 volume 0.37094E+03 ppm1 8.654 ppm2 3.049  
 (( segid "BRD" \* and resid 24 and name H4 ))  
 (( segid "BRD" \* and resid 24 and name H5 ))  
 ASST [ 1674 ] 2.800 2.000 2.000 peak 1691 weight 0.11000E+01 volume 0.12194E+04 ppm1 8.659 ppm2 9.112  
 (( segid "BRD" \* and resid 25 and name HA ))  
 (( segid "BRD" \* and resid 25 and name HN ))  
 ASST [ 1721 ] 3.200 1.900 peak 1711 weight 0.11000E+01 volume 0.81722E+03 ppm1 9.134 ppm2 8.155  
 (( segid "BRD" \* and resid 25 and name H1 ))  
 (( segid "BRD" \* and resid 25 and name H2 ))  
 ASST [ 1722 ] 3.200 2.400 2.300 peak 1721 weight 0.11000E+01 volume 0.16599E+03 ppm1 9.134 ppm2 4.443  
 (( segid "BRD" \* and resid 25 and name H3 ))  
 (( segid "BRD" \* and resid 25 and name H4 ))  
 ASST [ 1723 ] 3.200 2.400 2.300 peak 1731 weight 0.11000E+01 volume 0.41222E+03 ppm1 9.133 ppm2 1.799  
 (( segid "BRD" \* and resid 25 and name H5 ))  
 (( segid "BRD" \* and resid 25 and name H6 ))  
 ASST [ 1724 ] 3.200 2.400 2.300 peak 1741 weight 0.11000E+01 volume 0.35918E+03 ppm1 9.133 ppm2 1.627  
 (( segid "BRD" \* and resid 26 and name HA ))  
 (( segid "BRD" \* and resid 26 and name HN ))  
 ASST [ 1801 ] 3.000 2.200 2.200 peak 1791 weight 0.11000E+01 volume 0.34501E+03 ppm1 9.194 ppm2 4.489  
 (( segid "BRD" \* and resid 26 and name H1 ))  
 (( segid "BRD" \* and resid 26 and name H2 ))  
 ASST [ 1802 ] 3.000 2.200 2.200 peak 1801 weight 0.11000E+01 volume 0.50197E+03 ppm1 8.169 ppm2 2.480  
 (( segid "BRD" \* and resid 26 and name H3 ))  
 (( segid "BRD" \* and resid 26 and name H4 ))  
 ASST [ 1803 ] 3.000 2.200 2.200 peak 1811 weight 0.11000E+01 volume 0.49451E+03 ppm1 8.170 ppm2 5.055  
 (( segid "BRD" \* and resid 26 and name H5 ))  
 (( segid "BRD" \* and resid 26 and name H6 ))  
 ASST [ 1804 ] 3.000 2.200 2.200 peak 1821 weight 0.11000E+01 volume 0.78542E+03 ppm1 8.170 ppm2 4.411  
 (( segid "BRD" \* and resid 27 and name HA ))  
 (( segid "BRD" \* and resid 27 and name HN ))  
 ASST [ 1931 ] 2.800 1.700 1.700 peak 1921 weight 0.11000E+01 volume 0.56681E+03 ppm1 8.169 ppm2 9.179  
 (( segid "BRD" \* and resid 27 and name H1 ))  
 (( segid "BRD" \* and resid 27 and name H2 ))  
 ASST [ 1932 ] 2.800 1.700 1.700 peak 1931 weight 0.11000E+01 volume 0.20055E+03 ppm1 9.187 ppm2 4.645  
 (( segid "BRD" \* and resid 27 and name H3 ))  
 (( segid "BRD" \* and resid 27 and name H4 ))  
 ASST [ 1933 ] 2.800 2.000 2.000 peak 1941 weight 0.11000E+01 volume 0.37222E+03 ppm1 9.189 ppm2 4.285  
 (( segid "BRD" \* and resid 27 and name H5 ))  
 (( segid "BRD" \* and resid 27 and name H6 ))  
 ASST [ 1934 ] 2.800 2.000 2.000 peak 1951 weight 0.11000E+01 volume 0.33917E+03 ppm1 9.166 ppm2 9.170  
 (( segid "BRD" \* and resid 28 and name HA ))  
 (( segid "BRD" \* and resid 28 and name HN ))  
 ASST [ 1935 ] 2.800 2.000 2.000 peak 1961 weight 0.11000E+01 volume 0.35792E+03 ppm1 9.658 ppm2 4.447  
 (( segid "BRD" \* and resid 28 and name H1 ))  
 (( segid "BRD" \* and resid 28 and name H2 ))  
 ASST [ 1936 ] 2.800 2.000 2.000 peak 1971 weight 0.11000E+01 volume 0.43140E+03 ppm1 9.659 ppm2 4.917  
 (( segid "BRD" \* and resid 28 and name H3 ))  
 (( segid "BRD" \* and resid 28 and name H4 ))  
 ASST [ 1937 ] 2.800 2.000 2.000 peak 1981 weight 0.11000E+01 volume 0.24097E+03 ppm1 9.659 ppm2 1.910  
 (( segid "BRD" \* and resid 28 and name H5 ))  
 (( segid "BRD" \* and resid 28 and name H6 ))  
 ASST [ 1938 ] 2.800 2.000 2.000 peak 1991 weight 0.11000E+01 volume 0.37690E+03 ppm1 9.464 ppm2 4.901  
 (( segid "BRD" \* and resid 29 and name HA ))  
 (( segid "BRD" \* and resid 29 and name HN ))  
 ASST [ 1939 ] 2.800 2.000 2.000 peak 2001 weight 0.11000E+01 volume 0.42690E+03 ppm1 9.463 ppm2 3.592  
 (( segid "BRD" \* and resid 29 and name H1 ))  
 (( segid "BRD" \* and resid 29 and name H2 ))  
 ASST [ 1940 ] 2.800 2.000 2.000 peak 2011 weight 0.11000E+01 volume 0.26941E+03 ppm1 9.463 ppm2 3.268

ASST [ 1971 ]  
 (( segid "BRD" \* and resid 33 and name HN ))  
 (( segid "BRD" \* and resid 34 and name HN ))  
 ASST [ 1981 ] 1.700 1.700 peak 1971 weight 0.11000E+01 volume 0.41274E+03 ppm1 9.658 ppm2 9.450  
 (( segid "BRD" \* and resid 35 and name HN ))  
 (( segid "BRD" \* and resid 35 and name H1 ))  
 ASST [ 1991 ] 2.900 3.100 3.100 peak 1991 weight 0.11000E+01 volume 0.32481E+03 ppm1 7.516 ppm2 9.014  
 (( segid "BRD" \* and resid 35 and name H2 ))  
 (( segid "BRD" \* and resid 35 and name H3 ))  
 ASST [ 2001 ] 2.900 3.100 3.100 peak 2001 weight 0.11000E+01 volume 0.39507E+03 ppm1 7.516 ppm2 3.919  
 (( segid "BRD" \* and resid 36 and name HA ))  
 (( segid "BRD" \* and resid 36 and name HN ))  
 ASST [ 2011 ] 2.900 3.000 3.000 peak 2011 weight 0.11000E+01 volume 0.37475E+03 ppm1 9.464 ppm2 7.497  
 (( segid "BRD" \* and resid 36 and name H1 ))  
 (( segid "BRD" \* and resid 36 and name H2 ))  
 ASST [ 2021 ] 2.900 3.000 3.000 peak 2021 weight 0.11000E+01 volume 0.17051E+03 ppm1 8.423 ppm2 3.911  
 (( segid "BRD" \* and resid 36 and name H3 ))  
 (( segid "BRD" \* and resid 36 and name H4 ))  
 ASST [ 2031 ] 2.900 3.000 3.000 peak 2031 weight 0.11000E+01 volume 0.42952E+03 ppm1 8.423 ppm2 4.809  
 (( segid "BRD" \* and resid 37 and name HA ))  
 (( segid "BRD" \* and resid 37 and name HN ))  
 ASST [ 2041 ] 2.900 3.000 3.000 peak 2041 weight 0.11000E+01 volume 0.39745E+03 ppm1 7.516 ppm2 8.406  
 (( segid "BRD" \* and resid 37 and name H1 ))  
 (( segid "BRD" \* and resid 37 and name H2 ))  
 ASST [ 2051 ] 2.900 3.000 3.000 peak 2051 weight 0.11000E+01 volume 0.11243E+03 ppm1 8.570 ppm2 2.350  
 (( segid "BRD" \* and resid 37 and name H3 ))  
 (( segid "BRD" \* and resid 37 and name H4 ))  
 ASST [ 2061 ] 2.900 3.000 3.000 peak 2061 weight 0.11000E+01 volume 0.28579E+03 ppm1 8.570 ppm2 2.810  
 (( segid "BRD" \* and resid 37 and name H5 ))  
 (( segid "BRD" \* and resid 37 and name H6 ))  
 ASST [ 2071 ] 2.900 3.000 3.000 peak 2071 weight 0.11000E+01 volume 0.46735E+03 ppm1 8.571 ppm2 2.407  
 (( segid "BRD" \* and resid 38 and name HA ))  
 (( segid "BRD" \* and resid 38 and name HN ))  
 ASST [ 2081 ] 2.700 2.200 2.200 peak 2081 weight 0.11000E+01 volume 0.28244E+03 ppm1 8.572 ppm2 3.014  
 (( segid "BRD" \* and resid 38 and name H1 ))  
 (( segid "BRD" \* and resid 38 and name H2 ))  
 ASST [ 2091 ] 2.700 2.200 2.200 peak 2091 weight 0.11000E+01 volume 0.15320E+03 ppm1 8.355 ppm2 7.595  
 (( segid "BRD" \* and resid 38 and name H3 ))  
 (( segid "BRD" \* and resid 38 and name H4 ))  
 ASST [ 2101 ] 3.100 3.400 3.400 peak 2101 weight 0.11000E+01 volume 0.22276E+03 ppm1 8.356 ppm2 4.974  
 (( segid "BRD" \* and resid 38 and name H5 ))  
 (( segid "BRD" \* and resid 38 and name H6 ))  
 ASST [ 2111 ] 3.100 3.400 3.400 peak 2111 weight 0.11000E+01 volume 0.72297E+03 ppm1 8.354 ppm2 3.536  
 (( segid "BRD" \* and resid 39 and name HA ))  
 (( segid "BRD" \* and resid 39 and name HN ))  
 ASST [ 2121 ] 3.000 3.200 3.200 peak 2121 weight 0.11000E+01 volume 0.24047E+03 ppm1 8.858 ppm2 5.431  
 (( segid "BRD" \* and resid 39 and name H1 ))  
 (( segid "BRD" \* and resid 39 and name H2 ))  
 ASST [ 2131 ] 3.000 3.200 3.200 peak 2131 weight 0.11000E+01 volume 0.26996E+03 ppm1 8.858 ppm2 3.471  
 (( segid "BRD" \* and resid 39 and name H3 ))  
 (( segid "BRD" \* and resid 39 and name H4 ))  
 ASST [ 2141 ] 2.400 1.400 1.400 peak 2141 weight 0.11000E+01 volume 0.10141E+04 ppm1 8.858 ppm2 3.500  
 (( segid "BRD" \* and resid 39 and name H5 ))  
 (( segid "BRD" \* and resid 39 and name H6 ))  
 ASST [ 2201 ] 2.400 1.700 1.700 peak 2201 weight 0.11000E+01 volume 0.42485E+03 ppm1 8.868 ppm2 8.468  
 (( segid "BRD" \* and resid 46 and name HA ))  
 (( segid "BRD" \* and resid 46 and name HN ))  
 ASST [ 2211 ] 3.000 3.000 3.000 peak 2201 weight 0.11000E+01 volume 0.38079E+03 ppm1 8.542 ppm2 8.436  
 (( segid "BRD" \* and resid 46 and name H1 ))  
 (( segid "BRD" \* and resid 46 and name H2 ))  
 ASST [ 2221 ] 3.000 3.000 3.000 peak 2211 weight 0.11000E+01 volume 0.49246E+03 ppm1 8.542 ppm2 4.135  
 (( segid "BRD" \* and resid 46 and name H3 ))  
 (( segid "BRD" \* and resid 46 and name H4 ))  
 ASST [ 2231 ] 3.000 3.000 3.000 peak 2221 weight 0.11000E+01 volume 0.82312E+03 ppm1 8.561 ppm2 3.293  
 (( segid "BRD" \* and resid 46 and name H5 ))  
 (( segid "BRD" \* and resid 46 and name H6 ))  
 ASST [ 2241 ] 2.500 1.600 1.600 peak 2231 weight 0.11000E+01 volume 0.72608E+03 ppm1 8.842 ppm2 3.098





7

$\infty$ [illegible]

9

[illegible]





[illegible]

[illegible]

14



[illegible]



[illegible]





20







24



```

(( segid "BTD" and resid 10 and name HN ))
ASST ( 15551 ) 5.200 1.300 0.300 peak 15551 weight 0.10000E+01 volume 0.99878E+01 ppm1 8.485 ppm2 4.388 2.042 ppm2 0.759
(( segid "BTD" and resid 7 and name H2 ))
(( segid "BTD" and resid 15 and name H2 ))
(( segid "BTD" and resid 15 and name H2 ))
ASST ( 15551 ) 4.900 1.800 1.800 peak 15551 weight 0.10000E+01 volume 0.14338E+02 ppm1 8.924 ppm2 4.281 4.607 ppm2 3.811
(( segid "BTD" and resid 32 and name H2 ))
(( segid "BTD" and resid 33 and name H2 ))
(( segid "BTD" and resid 33 and name H2 ))
ASST ( 15551 ) 3.400 1.700 1.700 peak 15551 weight 0.10000E+01 volume 0.41447E+02 ppm1 11.082 ppm2 2.167 3.274 ppm2 4.143
(( segid "BTD" and resid 47 and name H2 ))
(( segid "BTD" and resid 47 and name H2 ))
(( segid "BTD" and resid 47 and name H2 ))
ASST ( 15551 ) 4.400 1.800 1.800 peak 15551 weight 0.10000E+01 volume 0.30408E+02 ppm1 8.832 ppm2 4.471 3.077 ppm2 4.143
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 3.200 1.900 1.900 peak 15551 weight 0.11000E+01 volume 0.39165E+02 ppm1 3.669 ppm2 4.556 4.704 ppm2 3.395
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 3.700 1.400 1.400 peak 15551 weight 0.11000E+01 volume 0.32058E+02 ppm1 3.472 ppm2 4.356 3.815 ppm2 4.721
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.800 1.700 1.700 peak 15551 weight 0.11000E+01 volume 0.44638E+02 ppm1 4.654 ppm2 3.572 4.654 ppm2 3.572
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.500 1.400 1.400 peak 15551 weight 0.11000E+01 volume 0.31584E+02 ppm1 4.453 ppm2 2.669 4.453 ppm2 2.669
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 3.100 2.400 2.400 peak 15551 weight 0.11000E+01 volume 0.94438E+02 ppm1 3.669 ppm2 5.143 3.669 ppm2 5.143
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.700 1.800 1.800 peak 15551 weight 0.11000E+01 volume 0.18471E+03 ppm1 3.522 ppm2 5.143 3.522 ppm2 5.143
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.400 1.700 1.700 peak 15551 weight 0.11000E+01 volume 0.28741E+03 ppm1 3.522 ppm2 4.949 3.522 ppm2 4.949
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.800 2.000 2.000 peak 15551 weight 0.11000E+01 volume 0.23932E+03 ppm1 4.455 ppm2 3.361 4.455 ppm2 3.361
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.400 1.700 1.700 peak 15551 weight 0.11000E+01 volume 0.28555E+03 ppm1 3.619 ppm2 4.444 3.619 ppm2 4.444
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.800 2.000 2.000 peak 15551 weight 0.11000E+01 volume 0.27185E+03 ppm1 4.409 ppm2 3.999 4.409 ppm2 3.999
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.500 1.400 1.400 peak 15551 weight 0.11000E+01 volume 0.16998E+03 ppm1 4.409 ppm2 3.116 4.409 ppm2 3.116
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.500 1.400 1.400 peak 15551 weight 0.11000E+01 volume 0.36348E+03 ppm1 5.000 ppm2 2.316 5.000 ppm2 2.316
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.300 1.300 1.300 peak 15551 weight 0.11000E+01 volume 0.74133E+03 ppm1 1.700 ppm2 5.539 1.700 ppm2 5.539
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.100 1.100 1.100 peak 15551 weight 0.11000E+01 volume 0.87848E+03 ppm1 2.093 ppm2 4.687 2.093 ppm2 4.687
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.000 1.000 1.000 peak 15551 weight 0.11000E+01 volume 0.10742E+04 ppm1 2.190 ppm2 4.440 2.190 ppm2 4.440
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.100 1.100 1.100 peak 15551 weight 0.11000E+01 volume 0.84248E+03 ppm1 4.901 ppm2 1.979 4.901 ppm2 1.979
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.400 1.700 1.700 peak 15551 weight 0.11000E+01 volume 0.27701E+03 ppm1 5.542 ppm2 4.107 5.542 ppm2 4.107
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 2.100 1.100 1.100 peak 15551 weight 0.11000E+01 volume 0.16932E+03 ppm1 5.542 ppm2 3.146 5.542 ppm2 3.146
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 3.000 2.200 2.200 peak 15551 weight 0.11000E+01 volume 0.10998E+03 ppm1 4.339 ppm2 2.863 4.339 ppm2 2.863
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
(( segid "BTD" and resid 99 and name H2 ))
ASST ( 15551 ) 3.100 2.400 2.400 peak 15551 weight 0.11000E+01 volume 0.88126E+03 ppm1 4.361 ppm2 3.000 4.361 ppm2 3.000

```





ASST ( 1642)  
 (( segid "BTD" and resid 50 and name HD11 ))  
 (( segid "BTD" and resid 50 and name HD11 ))  
 2.800 2.000 2.000 peak 1882 weight 0.11000E+01 volume 0.14998E+03 ppm1  
 1.154 ppm2 1.826 2.942 4.407 ppm2  
 ASST ( 1693)  
 (( segid "BTD" and resid 50 and name HD11 ))  
 (( segid "BTD" and resid 50 and name HD11 ))  
 2.300 1.300 1.300 peak 1693 weight 0.11000E+01 volume 0.49711E+03 ppm1  
 1.154 ppm2 1.408 2.644 ppm2  
 ASST ( 1703)  
 (( segid "BTD" and resid 50 and name HD11 ))  
 (( segid "BTD" and resid 50 and name HD11 ))  
 2.400 1.400 1.400 peak 1702 weight 0.11000E+01 volume 0.46662E+03 ppm1  
 1.154 ppm2 0.819 4.810 ppm2  
 ASST ( 1713)  
 (( segid "BTD" and resid 50 and name HD11 ))  
 (( segid "BTD" and resid 50 and name HD11 ))  
 2.400 1.400 1.400 peak 1712 weight 0.11000E+01 volume 0.45388E+03 ppm1  
 1.154 ppm2 0.594 4.799 2.816 ppm2  
 ASST ( 1743)  
 (( segid "BTD" and resid 50 and name HD21 ))  
 (( segid "BTD" and resid 50 and name HD21 ))  
 2.400 1.400 1.400 peak 1742 weight 0.11000E+01 volume 0.43008E+03 ppm1  
 1.006 ppm2 4.519 3.030 ppm2  
 ASST ( 1753)  
 (( segid "BTD" and resid 50 and name HD21 ))  
 (( segid "BTD" and resid 50 and name HD21 ))  
 2.300 1.300 1.300 peak 1752 weight 0.11000E+01 volume 0.41688E+03 ppm1  
 1.006 ppm2 4.874 2.782 ppm2  
 ASST ( 1773)  
 (( segid "BTD" and resid 50 and name HD21 ))  
 (( segid "BTD" and resid 50 and name HD21 ))  
 2.400 1.400 1.400 peak 1772 weight 0.11000E+01 volume 0.40990E+03 ppm1  
 1.006 ppm2 4.670 2.832 ppm2  
 ASST ( 1793)  
 (( segid "BTD" and resid 101 and name HA ))  
 (( segid "BTD" and resid 101 and name HA ))  
 2.400 1.400 1.400 peak 1792 weight 0.11000E+01 volume 0.20691E+03 ppm1  
 4.261 ppm2 1.578 2.537 ppm2  
 ASST ( 1813)  
 (( segid "BTD" and resid 101 and name HA ))  
 (( segid "BTD" and resid 101 and name HA ))  
 2.300 1.300 1.300 peak 1812 weight 0.11000E+01 volume 0.50350E+03 ppm1  
 1.596 ppm2 4.265 2.336 ppm2  
 ASST ( 1833)  
 (( segid "BTD" and resid 101 and name HA ))  
 (( segid "BTD" and resid 101 and name HA ))  
 2.300 1.300 1.300 peak 1832 weight 0.11000E+01 volume 0.60307E+03 ppm1  
 1.550 ppm2 1.806 2.467 ppm2  
 ASST ( 1853)  
 (( segid "BTD" and resid 101 and name HA ))  
 (( segid "BTD" and resid 101 and name HA ))  
 2.400 1.400 1.400 peak 1852 weight 0.11000E+01 volume 0.71179E+03 ppm1  
 1.550 ppm2 1.806 2.347 ppm2  
 ASST ( 1873)  
 (( segid "BTD" and resid 101 and name HA ))  
 (( segid "BTD" and resid 101 and name HA ))  
 2.400 1.400 1.400 peak 1872 weight 0.11000E+01 volume 0.84008E+02 ppm1  
 2.487 ppm2 1.588 3.670 ppm2  
 ASST ( 1893)  
 (( segid "BTD" and resid 101 and name HA ))  
 (( segid "BTD" and resid 101 and name HA ))  
 2.300 1.300 1.300 peak 1892 weight 0.11000E+01 volume 0.32478E+03 ppm1  
 1.648 ppm2 1.222 3.670 ppm2  
 ASST ( 1913)  
 (( segid "BTD" and resid 101 and name HA ))  
 (( segid "BTD" and resid 101 and name HA ))  
 2.400 1.400 1.400 peak 1912 weight 0.11000E+01 volume 0.24142E+03 ppm1  
 1.598 ppm2 2.357 2.779 ppm2  
 ASST ( 1933)  
 (( segid "BTD" and resid 101 and name HA ))  
 (( segid "BTD" and resid 101 and name HA ))  
 2.300 1.300 1.300 peak 1932 weight 0.11000E+01 volume 0.23587E+03 ppm1  
 1.206 ppm2 2.508 4.901 ppm2  
 ASST ( 1953)  
 (( segid "BTD" and resid 101 and name HA ))  
 (( segid "BTD" and resid 101 and name HA ))  
 2.400 1.400 1.400 peak 1952 weight 0.11000E+01 volume 0.44094E+03 ppm1  
 1.205 ppm2 2.355 4.262 ppm2  
 ASST ( 1973)  
 (( segid "BTD" and resid 101 and name HA ))  
 (( segid "BTD" and resid 101 and name HA ))  
 2.300 1.300 1.300 peak 1972 weight 0.11000E+01 volume 0.49470E+03 ppm1  
 1.205 ppm2 1.588 4.280 ppm2  
 ASST ( 1993)  
 (( segid "BTD" and resid 112 and name HA ))  
 (( segid "BTD" and resid 112 and name HA ))  
 2.300 1.300 1.300 peak 1992 weight 0.11000E+01 volume 0.28568E+03 ppm1  
 1.993 ppm2 1.325 4.901 ppm2







ASB1 ( 4942 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 4942 weight 0.11000E+01 volume 0.29632E+03 ppm1 2.141 ppm2 1.417 4.686

ASB1 ( 4973 )  
 (( egid "grd" - and resid 109 and name HB1 ))  
 (( egid "grd" - and resid 109 and name HB1 ))  
 2.500 1.600 1.600 peak 4973 weight 0.11000E+01 volume 0.13461E+03 ppm1 2.487 ppm2 4.687

ASB1 ( 4983 )  
 (( egid "grd" - and resid 109 and name HB1 ))  
 (( egid "grd" - and resid 109 and name HB1 ))  
 2.500 1.600 1.600 peak 4983 weight 0.11000E+01 volume 0.54731E+03 ppm1 4.459 ppm2 1.796

ASB1 ( 5001 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5001 weight 0.11000E+01 volume 0.34838E+03 ppm1 4.459 ppm2 1.946

ASB1 ( 5013 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5013 weight 0.11000E+01 volume 0.20757E+03 ppm1 1.945 ppm2 4.468

ASB1 ( 5023 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5023 weight 0.11000E+01 volume 0.37628E+03 ppm1 1.795 ppm2 4.468

ASB1 ( 5033 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5033 weight 0.11000E+01 volume 0.14652E+03 ppm1 1.946 ppm2 3.606

ASB1 ( 5043 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5043 weight 0.11000E+01 volume 0.40668E+03 ppm1 1.796 ppm2 3.606

ASB1 ( 5053 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5053 weight 0.11000E+01 volume 0.34659E+03 ppm1 3.577 ppm2 1.945

ASB1 ( 5063 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5063 weight 0.11000E+01 volume 0.34617E+03 ppm1 3.577 ppm2 1.795

ASB1 ( 5073 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5073 weight 0.11000E+01 volume 0.20415E+03 ppm1 3.578 ppm2 4.463

ASB1 ( 5083 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5083 weight 0.11000E+01 volume 0.31370E+03 ppm1 5.592 ppm2 3.686

ASB1 ( 5103 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5103 weight 0.11000E+01 volume 0.16779E+03 ppm1 5.592 ppm2 3.533

ASB1 ( 5123 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5123 weight 0.11000E+01 volume 0.19830E+03 ppm1 3.228 ppm2 4.517

ASB1 ( 5133 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5133 weight 0.11000E+01 volume 0.27353E+03 ppm1 3.227 ppm2 2.847

ASB1 ( 5143 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5143 weight 0.11000E+01 volume 0.46028E+03 ppm1 3.533 ppm2 2.847

ASB1 ( 5173 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5173 weight 0.11000E+01 volume 0.31566E+03 ppm1 3.533 ppm2 2.931

ASB1 ( 5193 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5193 weight 0.11000E+01 volume 0.21836E+03 ppm1 3.227 ppm2 2.931

ASB1 ( 5203 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5203 weight 0.11000E+01 volume 0.16215E+03 ppm1 4.509 ppm2 3.520

ASB1 ( 5213 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5213 weight 0.11000E+01 volume 0.12463E+03 ppm1 1.989 ppm2 4.807

ASB1 ( 5223 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5223 weight 0.11000E+01 volume 0.27005E+03 ppm1 1.989 ppm2 4.397

ASB1 ( 5233 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5233 weight 0.11000E+01 volume 0.10390E+03 ppm1 5.543 ppm2 1.944

ASB1 ( 5273 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5273 weight 0.11000E+01 volume 0.89029E+03 ppm1 2.436 ppm2 3.572

ASB1 ( 5323 )  
 (( egid "grd" - and resid 109 and name HB2 ))  
 (( egid "grd" - and resid 109 and name HB2 ))  
 2.500 1.600 1.600 peak 5323 weight 0.11000E+01 volume 0.39102E+03 ppm1 3.137 ppm2 4.911

33



















[illegible]



43





46

```

(( segid "BTD" - and resid 59 and name H21 ))
ASST [13092]
3.500 3.100 2.000 peak 13092 weight 0.10000E+01 volume 0.44884E+02 ppm1 3.177 ppm2 2.484 4.360 ppm2 3.499
(( segid "BTD" - and resid 54 and name H21 ))
(( segid "BTD" - and resid 61 and name H21 ))
ASST [13113]
2.900 2.100 peak 13092 weight 0.10000E+01 volume 0.51666E+02 ppm1 3.288 ppm2 1.075 2.091 ppm2 2.344
(( segid "BTD" - and resid 71 and name H21 ))
(( segid "BTD" - and resid 64 and name H21 ))
ASST [13123]
3.000 2.200 2.200 peak 13112 weight 0.10000E+01 volume 0.10740E+03 ppm1 2.486 ppm2 5.142 1.402 ppm2 5.362
(( segid "BTD" - and resid 71 and name H21 ))
(( segid "BTD" - and resid 64 and name H21 ))
ASST [13123]
2.900 2.100 peak 13112 weight 0.10000E+01 volume 0.13246E+03 ppm1 2.583 ppm2 5.141 4.655 ppm2 1.987
(( segid "BTD" - and resid 64 and name H21 ))
(( segid "BTD" - and resid 54 and name H21 ))
ASST [13123]
3.100 2.000 peak 13122 weight 0.10000E+01 volume 0.45934E+02 ppm1 3.275 ppm2 0.790 3.667 ppm2 4.933
(( segid "BTD" - and resid 54 and name H21 ))
(( segid "BTD" - and resid 64 and name H21 ))
OR [13123]
3.100 2.000 2.400 peak 13122 weight 0.10000E+01 volume 0.84018E+02 ppm1 4.607 ppm2 1.637 4.459 ppm2 2.136
(( segid "BTD" - and resid 54 and name H21 ))
(( segid "BTD" - and resid 64 and name H21 ))
ASST [13123]
2.900 2.100 peak 13112 weight 0.10000E+01 volume 0.13246E+03 ppm1 2.583 ppm2 5.141 4.457 ppm2 3.581
(( segid "BTD" - and resid 64 and name H21 ))
(( segid "BTD" - and resid 54 and name H21 ))
ASST [13123]
2.700 1.800 1.800 peak 13112 weight 0.10000E+01 volume 0.18778E+03 ppm1 1.546 ppm2 1.145 9.000 ppm2 2.218
(( segid "BTD" - and resid 22 and name H21 ))
(( segid "BTD" - and resid 63 and name H21 ))
ASST [13123]
2.500 1.600 1.400 peak 13112 weight 0.10000E+01 volume 0.14390E+03 ppm1 1.648 ppm2 2.516 2.289 ppm2 4.437
(( segid "BTD" - and resid 22 and name H21 ))
(( segid "BTD" - and resid 63 and name H21 ))
ASST [13123]
2.900 2.100 peak 13112 weight 0.10000E+01 volume 0.13246E+03 ppm1 2.583 ppm2 5.141 2.338 ppm2 1.318
(( segid "BTD" - and resid 49 and name H21 ))
(( segid "BTD" - and resid 50 and name H21 ))
ASST [13172]
3.500 3.500 0.000 peak 13172 weight 0.10000E+01 volume 0.10586E+03 ppm1 2.289 ppm2 4.437 1.499 ppm2 2.495
(( segid "BTD" - and resid 49 and name H21 ))
(( segid "BTD" - and resid 50 and name H21 ))
ASST [13172]
3.100 2.000 peak 13172 weight 0.10000E+01 volume 0.10586E+03 ppm1 2.289 ppm2 4.437 3.769 ppm2 1.830
(( segid "BTD" - and resid 43 and name H21 ))
(( segid "BTD" - and resid 43 and name H21 ))
ASST [13172]
2.700 1.800 1.800 peak 13172 weight 0.10000E+01 volume 0.18778E+03 ppm1 1.546 ppm2 1.145 4.555 ppm2 1.435
(( segid "BTD" - and resid 43 and name H21 ))
(( segid "BTD" - and resid 43 and name H21 ))
ASST [13172]
3.500 3.500 0.000 peak 13172 weight 0.10000E+01 volume 0.10586E+03 ppm1 2.289 ppm2 4.437 4.411 ppm2 2.181
(( segid "BTD" - and resid 43 and name H21 ))
(( segid "BTD" - and resid 43 and name H21 ))
OR [14442]
2.800 2.000 3.000 peak 14442 weight 0.10000E+01 volume 0.16018E+03 ppm1 4.411 ppm2 2.181 1.254 ppm2 1.330
(( segid "BTD" - and resid 111 and name H21 ))
(( segid "BTD" - and resid 111 and name H21 ))
ASST [14442]
3.500 3.500 0.000 peak 14442 weight 0.10000E+01 volume 0.10586E+03 ppm1 2.289 ppm2 4.437 1.389 ppm2 1.278
(( segid "BTD" - and resid 111 and name H21 ))
(( segid "BTD" - and resid 111 and name H21 ))
ASST [14442]
3.100 2.000 1.900 peak 14442 weight 0.10000E+01 volume 0.10586E+03 ppm1 2.289 ppm2 4.437 3.864 ppm2 2.331
(( segid "BTD" - and resid 111 and name H21 ))
(( segid "BTD" - and resid 111 and name H21 ))
ASST [14442]
2.700 1.800 1.800 peak 14442 weight 0.10000E+01 volume 0.10586E+03 ppm1 2.289 ppm2 4.437 1.057 ppm2 2.467
(( segid "BTD" - and resid 21 and name H21 ))
(( segid "BTD" - and resid 21 and name H21 ))
ASST [14442]
2.900 2.400 2.400 peak 14442 weight 0.10000E+01 volume 0.10586E+03 ppm1 2.289 ppm2 4.437 1.205 ppm2 1.488
(( segid "BTD" - and resid 21 and name H21 ))
(( segid "BTD" - and resid 21 and name H21 ))
ASST [14442]
2.700 1.800 1.800 peak 14442 weight 0.10000E+01 volume 0.10586E+03 ppm1 2.289 ppm2 4.437 2.539 ppm2 1.075
(( segid "BTD" - and resid 51 and name H21 ))
(( segid "BTD" - and resid 51 and name H21 ))
ASST [14442]
3.500 3.500 0.000 peak 14442 weight 0.10000E+01 volume 0.10586E+03 ppm1 2.289 ppm2 4.437 4.460 ppm2 4.222
(( segid "BTD" - and resid 51 and name H21 ))
(( segid "BTD" - and resid 51 and name H21 ))
OR [15012]
2.300 1.300 1.300 peak 15012 weight 0.10000E+01 volume 0.40036E+03 ppm1 5.446 ppm2 2.986

```













53

ASFI [24832] 3.300 2.700 2.200 peak 24782 weight 0.10000E+01 volume 0.51590E+02 ppm1 1.495 ppm2 7.487 4.804 ppm2 5.444  
 (( segid "BTD" and resid 19 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.37891E+03 ppm1 2.191 ppm2 7.469 4.804 ppm2 7.779  
 (( segid "BTD" and resid 19 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 2.200 2.300 peak 24832 weight 0.10000E+01 volume 0.74438E+03 ppm1 2.191 ppm2 1.454 4.013 ppm2 4.933  
 (( segid "BTD" and resid 19 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.700 2.700 peak 24832 weight 0.10000E+01 volume 0.24553E+03 ppm1 3.124 ppm2 1.888 4.013 ppm2 4.933  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.600 2.400 peak 24832 weight 0.10000E+01 volume 0.25958E+03 ppm1 3.066 ppm2 1.888 4.013 ppm2 4.933  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.41203E+03 ppm1 2.486 ppm2 7.519 3.420 ppm2 2.206  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.15108E+03 ppm1 3.914 ppm2 2.206 3.914 ppm2 2.206  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.12121E+03 ppm1 3.470 ppm2 7.893 3.470 ppm2 7.893  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.45042E+03 ppm1 3.420 ppm2 7.779 3.420 ppm2 7.779  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.22222E+03 ppm1 3.917 ppm2 7.779 3.917 ppm2 7.779  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.41073E+03 ppm1 1.551 ppm2 4.532 1.551 ppm2 4.532  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.10658E+03 ppm1 3.949 ppm2 4.932 3.949 ppm2 4.932  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.38225E+02 ppm1 1.559 ppm2 0.408 1.559 ppm2 0.408  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.38428E+02 ppm1 1.559 ppm2 3.069 1.559 ppm2 3.069  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.20266E+02 ppm1 1.559 ppm2 3.012 1.559 ppm2 3.012  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.18743E+03 ppm1 1.559 ppm2 2.157 1.559 ppm2 2.157  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.84468E+02 ppm1 1.549 ppm2 1.970 1.549 ppm2 1.970  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.55301E+03 ppm1 2.338 ppm2 1.410 2.338 ppm2 1.410  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.19966E+02 ppm1 1.497 ppm2 0.476 1.497 ppm2 0.476  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.84474E+02 ppm1 1.154 ppm2 4.511 1.154 ppm2 4.511  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.14960E+02 ppm1 1.254 ppm2 4.569 1.254 ppm2 4.569  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.71378E+02 ppm1 1.261 ppm2 7.769 1.261 ppm2 7.769  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.39658E+03 ppm1 1.155 ppm2 7.438 1.155 ppm2 7.438  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.22034E+03 ppm1 1.155 ppm2 7.529 1.155 ppm2 7.529  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))  
 ASFI [24832] 2.400 1.400 2.400 peak 24832 weight 0.10000E+01 volume 0.31339E+03 ppm1 1.184 ppm2 2.182 1.184 ppm2 2.182  
 (( segid "BTD" and resid 21 and name H21 ))  
 (( segid "BTD" and resid 19 and name H21 ))



```

(( segid "BRD" - and resid 36 and name HA ))
ASST [27440] 2.900 2.100 peak 27392 weight 0.10000E+01 volume 0.54954E+02 ppm1 2.979 ppm2 5.442
(( segid "BRD" - and resid 69 and name HD21 ))
(( segid "BRD" - and resid 14 and name HD11 ))
(( segid "BRD" - and resid 18 and name HD11 ))
ASST [27451] 2.400 2.100 peak 27642 weight 0.10000E+01 volume 0.37753E+02 ppm1 1.547 ppm2 3.091
(( segid "BRD" - and resid 22 and name HD21 ))
(( segid "BRD" - and resid 25 and name HD11 ))
(( segid "BRD" - and resid 35 and name HD11 ))
ASST [27460] 3.400 1.400 peak 27452 weight 0.10000E+01 volume 0.30318E+02 ppm1 1.599 ppm2 1.788
(( segid "BRD" - and resid 32 and name HD11 ))
(( segid "BRD" - and resid 35 and name HD11 ))
(( segid "BRD" - and resid 38 and name HD11 ))
ASST [27492] 3.400 2.100 peak 27472 weight 0.10000E+01 volume 0.54053E+02 ppm1 1.599 ppm2 3.003
(( segid "BRD" - and resid 60 and name HA ))
(( segid "BRD" - and resid 69 and name HD21 ))
(( segid "BRD" - and resid 74 and name HD11 ))
ASST [27552] 3.300 1.100 peak 27652 weight 0.10000E+01 volume 0.56630E+02 ppm1 1.599 ppm2 4.409
(( segid "BRD" - and resid 74 and name HD11 ))
(( segid "BRD" - and resid 74 and name HD11 ))
(( segid "BRD" - and resid 74 and name HD11 ))
ASST [27579] 2.300 2.300 peak 27552 weight 0.10000E+01 volume 0.54070E+02 ppm1 1.599 ppm2 7.529
(( segid "BRD" - and resid 22 and name HD11 ))
(( segid "BRD" - and resid 22 and name HD11 ))
(( segid "BRD" - and resid 22 and name HD11 ))
ASST [27642] 2.400 2.300 peak 27572 weight 0.10000E+01 volume 0.60681E+02 ppm1 1.645 ppm2 4.990
(( segid "BRD" - and resid 22 and name HD11 ))
(( segid "BRD" - and resid 22 and name HD11 ))
(( segid "BRD" - and resid 22 and name HD11 ))
ASST [27642] 2.500 1.400 peak 27632 weight 0.10000E+01 volume 0.31070E+02 ppm1 1.549 ppm2 5.143
(( segid "BRD" - and resid 73 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
ASST [27692] 2.400 2.400 peak 27662 weight 0.10000E+01 volume 0.86052E+02 ppm1 1.549 ppm2 4.362
(( segid "BRD" - and resid 73 and name HD11 ))
(( segid "BRD" - and resid 73 and name HD11 ))
(( segid "BRD" - and resid 73 and name HD11 ))
ASST [27712] 2.400 2.000 peak 27692 weight 0.10000E+01 volume 0.17686E+02 ppm1 1.500 ppm2 3.104
(( segid "BRD" - and resid 35 and name HA ))
(( segid "BRD" - and resid 35 and name HA ))
(( segid "BRD" - and resid 35 and name HA ))
ASST [27722] 3.400 2.100 peak 27712 weight 0.10000E+01 volume 0.10158E+02 ppm1 4.900 ppm2 2.004
(( segid "BRD" - and resid 48 and name HA ))
(( segid "BRD" - and resid 48 and name HA ))
(( segid "BRD" - and resid 48 and name HA ))
ASST [27772] 3.000 2.200 peak 27722 weight 0.10000E+01 volume 0.11137E+02 ppm1 4.803 ppm2 1.652
(( segid "BRD" - and resid 70 and name HD11 ))
(( segid "BRD" - and resid 70 and name HD11 ))
(( segid "BRD" - and resid 70 and name HD11 ))
ASST [27822] 3.400 2.900 2.100 peak 27772 weight 0.10000E+01 volume 0.55089E+02 ppm1 2.487 ppm2 4.377
(( segid "BRD" - and resid 76 and name HA ))
(( segid "BRD" - and resid 77 and name HA ))
(( segid "BRD" - and resid 77 and name HA ))
ASST [27842] 3.400 1.700 peak 27822 weight 0.10000E+01 volume 0.26285E+02 ppm1 3.947 ppm2 3.312
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
ASST [27872] 3.700 3.400 1.800 peak 27842 weight 0.10000E+01 volume 0.31945E+02 ppm1 1.056 ppm2 1.954
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
ASST [27892] 3.400 2.100 peak 27892 weight 0.10000E+01 volume 0.17221E+02 ppm1 1.056 ppm2 4.411
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
ASST [27922] 5.500 5.500 0.000 peak 27922 weight 0.10000E+01 volume 0.22274E+02 ppm1 1.254 ppm2 3.003
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
OR [27992] 2.400 2.000 2.000 peak 27992 weight 0.10000E+01 volume 0.30077E+02 ppm1 1.254 ppm2 3.451
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
ASST [27992] 2.400 1.700 1.700 peak 27942 weight 0.10000E+01 volume 0.24618E+02 ppm1 0.740 ppm2 7.435
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
ASST [27992] 2.300 1.200 peak 27992 weight 0.10000E+01 volume 0.74333E+02 ppm1 0.642 ppm2 7.526
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
(( segid "BRD" - and resid 76 and name HD11 ))
OR [27992] 3.100 1.100 peak 27992 weight 0.10000E+01 volume 0.63476E+02 ppm1 0.642 ppm2 7.031

```







59

3.700 3.400 1.800 peak 1753 weight 0.10000E+01 volume 0.11649E+02 ppm1 7.617 ppm2 4.534  
 ASSE ( 180 )  
 (( segid "BPD" - and resid 85 and name HD1 ))  
 (( segid "BPD" - and resid 85 and name HD1 ))  
 3.300 2.700 2.200 peak 1803 weight 0.10000E+01 volume 0.27238E+02 ppm1 7.476 ppm2 4.985  
 ASSE ( 183 )  
 (( segid "BPD" - and resid 95 and name HD1 ))  
 (( segid "BPD" - and resid 95 and name HD1 ))  
 3.700 3.700 2.700 2.200 peak 1853 weight 0.10000E+01 volume 0.27537E+02 ppm1 7.479 ppm2 3.920  
 ASSE ( 190 )  
 (( segid "BPD" - and resid 95 and name HD1 ))  
 (( segid "BPD" - and resid 95 and name HD1 ))  
 3.200 2.600 2.300 peak 1863 weight 0.10000E+01 volume 0.30308E+02 ppm1 7.476 ppm2 1.047  
 ASSE ( 193 )  
 (( segid "BPD" - and resid 95 and name HD1 ))  
 (( segid "BPD" - and resid 95 and name HD1 ))  
 3.100 3.100 2.000 peak 1903 weight 0.10000E+01 volume 0.17945E+02 ppm1 7.616 ppm2 4.343  
 ASSE ( 200 )  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 4.300 4.300 1.300 peak 2003 weight 0.10000E+01 volume 0.52192E+01 ppm1 7.757 ppm2 7.509  
 ASSE ( 214 )  
 (( segid "BPD" - and resid 42 and name HD1 ))  
 (( segid "BPD" - and resid 42 and name HD1 ))  
 3.800 3.400 1.700 peak 2143 weight 0.10000E+01 volume 0.11137E+02 ppm1 7.013 ppm2 7.509  
 ASSE ( 218 )  
 (( segid "BPD" - and resid 107 and name HD1 ))  
 (( segid "BPD" - and resid 107 and name HD1 ))  
 2.500 3.100 2.000 peak 2183 weight 0.10000E+01 volume 0.13605E+02 ppm1 7.897 ppm2 3.060  
 ASSE ( 230 )  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 2.200 1.200 1.200 peak 2303 weight 0.10000E+01 volume 0.93346E+03 ppm1 7.970 ppm2 3.806  
 ASSE ( 244 )  
 (( segid "BPD" - and resid 47 and name HD1 ))  
 (( segid "BPD" - and resid 47 and name HD1 ))  
 2.700 2.700 1.300 peak 2443 weight 0.10000E+01 volume 0.78671E+03 ppm1 7.970 ppm2 3.415  
 ASSE ( 274 )  
 (( segid "BPD" - and resid 67 and name HD1 ))  
 (( segid "BPD" - and resid 67 and name HD1 ))  
 2.000 1.000 1.000 peak 2743 weight 0.10000E+01 volume 0.18744E+04 ppm1 6.900 ppm2 7.321  
 ASSE ( 314 )  
 (( segid "BPD" - and resid 67 and name HD1 ))  
 (( segid "BPD" - and resid 67 and name HD1 ))  
 2.300 1.300 1.300 peak 3143 weight 0.10000E+01 volume 0.81340E+03 ppm1 7.776 ppm2 3.546  
 ASSE ( 354 )  
 (( segid "BPD" - and resid 68 and name HD1 ))  
 (( segid "BPD" - and resid 68 and name HD1 ))  
 2.400 2.400 1.000 peak 3543 weight 0.10000E+01 volume 0.17752E+04 ppm1 7.612 ppm2 7.416  
 ASSE ( 454 )  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 3.000 2.200 2.200 peak 4543 weight 0.10000E+01 volume 0.17621E+03 ppm1 8.493 ppm2 4.977  
 ASSE ( 474 )  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 3.400 3.400 2.400 peak 4743 weight 0.10000E+01 volume 0.14234E+03 ppm1 8.490 ppm2 4.213  
 ASSE ( 1594 )  
 (( segid "BPD" - and resid 34 and name HD1 ))  
 (( segid "BPD" - and resid 34 and name HD1 ))  
 3.300 2.700 2.300 peak 15943 weight 0.10000E+01 volume 0.91431E+02 ppm1 7.781 ppm2 4.116  
 ASSE ( 1634 )  
 (( segid "BPD" - and resid 96 and name HD1 ))  
 (( segid "BPD" - and resid 96 and name HD1 ))  
 3.800 3.800 0.900 peak 16343 weight 0.10000E+01 volume 0.28744E+04 ppm1 7.774 ppm2 7.614  
 ASSE ( 1714 )  
 (( segid "BPD" - and resid 96 and name HD1 ))  
 (( segid "BPD" - and resid 96 and name HD1 ))  
 2.200 1.200 1.200 peak 17143 weight 0.10000E+01 volume 0.99570E+03 ppm1 7.736 ppm2 3.124  
 ASSE ( 2024 )  
 (( segid "BPD" - and resid 34 and name HD1 ))  
 (( segid "BPD" - and resid 34 and name HD1 ))  
 2.500 1.400 1.400 peak 20243 weight 0.10000E+01 volume 0.48715E+03 ppm1 7.714 ppm2 4.115  
 ASSE ( 2084 )  
 (( segid "BPD" - and resid 15 and name HD1 ))  
 (( segid "BPD" - and resid 15 and name HD1 ))  
 1.900 0.900 0.900 peak 20843 weight 0.10000E+01 volume 0.28924E+04 ppm1 7.689 ppm2 7.483  
 ASSE ( 2474 )  
 (( segid "BPD" - and resid 96 and name HD1 ))  
 (( segid "BPD" - and resid 96 and name HD1 ))  
 3.300 2.700 2.200 peak 24743 weight 0.10000E+01 volume 0.86609E+02 ppm1 7.611 ppm2 4.001  
 ASSE ( 2534 )  
 (( segid "BPD" - and resid 74 and name HD1 ))  
 (( segid "BPD" - and resid 74 and name HD1 ))  
 1.000 1.000 peak 25343 weight 0.10000E+01 volume 0.14278E+04 ppm1 7.539 ppm2 7.006  
 ASSE ( 3204 )

(( segid "BPD" - and resid 82 and name HD1 ))  
 (( segid "BPD" - and resid 82 and name HD1 ))  
 2.400 1.400 1.400 peak 32043 weight 0.10000E+01 volume 0.67786E+03 ppm1 7.265 ppm2 3.708  
 ASSE ( 3214 )  
 (( segid "BPD" - and resid 82 and name HD1 ))  
 (( segid "BPD" - and resid 82 and name HD1 ))  
 2.300 1.300 1.300 peak 32143 weight 0.10000E+01 volume 0.77298E+03 ppm1 7.262 ppm2 3.571  
 ASSE ( 3304 )  
 (( segid "BPD" - and resid 82 and name HD1 ))  
 (( segid "BPD" - and resid 82 and name HD1 ))  
 2.100 1.100 1.100 peak 33043 weight 0.10000E+01 volume 0.12643E+04 ppm1 7.262 ppm2 7.076  
 ASSE ( 3444 )  
 (( segid "BPD" - and resid 74 and name HD1 ))  
 (( segid "BPD" - and resid 74 and name HD1 ))  
 2.400 1.400 1.400 peak 34443 weight 0.10000E+01 volume 0.64123E+03 ppm1 7.005 ppm2 3.007  
 ASSE ( 3894 )  
 (( segid "BPD" - and resid 46 and name HD1 ))  
 (( segid "BPD" - and resid 46 and name HD1 ))  
 2.000 1.000 1.000 peak 38943 weight 0.10000E+01 volume 0.17599E+04 ppm1 5.743 ppm2 6.687  
 ASSE ( 4044 )  
 (( segid "BPD" - and resid 46 and name HD1 ))  
 (( segid "BPD" - and resid 46 and name HD1 ))  
 2.400 1.400 1.400 peak 39143 weight 0.10000E+01 volume 0.61947E+03 ppm1 5.740 ppm2 3.304  
 ASSE ( 4284 )  
 (( segid "BPD" - and resid 28 and name HD1 ))  
 (( segid "BPD" - and resid 28 and name HD1 ))  
 2.000 2.000 2.200 peak 42843 weight 0.10000E+01 volume 0.15444E+03 ppm1 5.577 ppm2 4.586  
 ASSE ( 484 )  
 (( segid "BPD" - and resid 82 and name HD1 ))  
 (( segid "BPD" - and resid 103 and name HD1 ))  
 2.900 2.100 2.100 peak 843 weight 0.10000E+01 volume 0.22432E+03 ppm1 7.265 ppm2 3.790  
 ASSE ( 174 )  
 (( segid "BPD" - and resid 67 and name HD1 ))  
 (( segid "BPD" - and resid 67 and name HD1 ))  
 3.300 2.700 2.200 peak 1743 weight 0.10000E+01 volume 0.89709E+02 ppm1 6.899 ppm2 4.490  
 ASSE ( 244 )  
 (( segid "BPD" - and resid 67 and name HD1 ))  
 (( segid "BPD" - and resid 67 and name HD1 ))  
 2.900 2.100 2.100 peak 2443 weight 0.10000E+01 volume 0.21246E+03 ppm1 6.899 ppm2 1.706  
 ASSE ( 314 )  
 (( segid "BPD" - and resid 67 and name HD1 ))  
 (( segid "BPD" - and resid 73 and name HD1 ))  
 3.500 3.100 2.000 peak 2143 weight 0.10000E+01 volume 0.64743E+02 ppm1 6.899 ppm2 1.545  
 ASSE ( 454 )  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 (( segid "BPD" - and resid 29 and name HD1 ))  
 2.700 2.700 2.200 peak 4543 weight 0.10000E+01 volume 0.97318E+02 ppm1 8.490 ppm2 4.814  
 ASSE ( 494 )  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 5.500 5.500 0.000 peak 4943 weight 0.10000E+01 volume 0.38800E+03 ppm1 8.490 ppm2 2.781  
 ASSE ( 504 )  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 4.000 4.000 1.800 peak 5043 weight 0.10000E+01 volume 0.30117E+02 ppm1 8.490 ppm2 3.211  
 ASSE ( 544 )  
 (( segid "BPD" - and resid 107 and name HD1 ))  
 (( segid "BPD" - and resid 79 and name HD1 ))  
 3.100 2.400 2.400 peak 5443 weight 0.10000E+01 volume 0.14239E+03 ppm1 8.008 ppm2 2.802  
 ASSE ( 624 )  
 (( segid "BPD" - and resid 107 and name HD1 ))  
 (( segid "BPD" - and resid 71 and name HD1 ))  
 3.000 2.200 2.200 peak 6243 weight 0.10000E+01 volume 0.17938E+03 ppm1 8.058 ppm2 4.439  
 ASSE ( 634 )  
 (( segid "BPD" - and resid 107 and name HD1 ))  
 (( segid "BPD" - and resid 107 and name HD1 ))  
 2.000 1.000 1.000 peak 6343 weight 0.10000E+01 volume 0.19563E+04 ppm1 8.058 ppm2 7.927  
 ASSE ( 644 )  
 (( segid "BPD" - and resid 107 and name HD1 ))  
 (( segid "BPD" - and resid 79 and name HD1 ))  
 2.500 2.500 2.000 peak 6443 weight 0.10000E+01 volume 0.49678E+03 ppm1 8.058 ppm2 7.793  
 ASSE ( 714 )  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 3.100 2.100 2.400 peak 7143 weight 0.10000E+01 volume 0.11157E+04 ppm1 8.004 ppm2 7.797  
 ASSE ( 724 )  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 (( segid "BPD" - and resid 33 and name HD1 ))  
 3.200 2.600 2.300 peak 7243 weight 0.10000E+01 volume 0.11962E+03 ppm1 7.959 ppm2 -0.311  
 ASSE ( 804 )  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 2.700 1.800 1.800 peak 8043 weight 0.10000E+01 volume 0.23995E+02 ppm1 7.961 ppm2 2.667  
 ASSE ( 854 )  
 (( segid "BPD" - and resid 32 and name HD1 ))  
 (( segid "BPD" - and resid 32 and name HD1 ))

19

62

63

### Ambiguous NOE-derived Inter-proton Distance Restraints

[illegible]







4

5



7





```

(( segid "BTD" and resid 14 and name HD11)
OR (( segid "BTD" and resid 17 and name HB ))
(( segid "BTD" and resid 19 and name HD1 ))
ASST (( 7232)
(( segid "BTD" and resid 17 and name HB ))
(( segid "BTD" and resid 115 and name HD11)
3.400 2.900 2.100 peak 7232 weight
OR (( segid "BTD" and resid 17 and name HB ))
(( segid "BTD" and resid 14 and name HD21)
ASST (( 7243)
(( segid "BTD" and resid 17 and name HD21)
(( segid "BTD" and resid 31 and name HD11)
OR (( 7242)
3.400 1.400 1.400 peak 7242 weight
(( segid "BTD" and resid 17 and name HD21)
ASST (( 7192)
(( segid "BTD" and resid 15 and name HA ))
(( segid "BTD" and resid 18 and name HG ))
OR (( 7193)
3.400 1.700 peak 7192 weight
(( segid "BTD" and resid 15 and name HA ))
(( segid "BTD" and resid 19 and name HB1 ))
ASST (( 7743)
(( segid "BTD" and resid 106 and name HA ))
(( segid "BTD" and resid 35 and name HD11)
OR (( 7742)
3.400 2.700 2.100 peak 7742 weight
(( segid "BTD" and resid 104 and name HA ))
(( segid "BTD" and resid 102 and name HB2 ))
ASST (( 8052)
(( segid "BTD" and resid 77 and name HB1 ))
(( segid "BTD" and resid 80 and name HB2 ))
OR (( 8051)
3.400 2.700 2.100 peak 8052 weight
(( segid "BTD" and resid 77 and name HB1 ))
OR (( 1052)
(( segid "BTD" and resid 77 and name HB1 ))
(( segid "BTD" and resid 54 and name HB1 ))
(( segid "BTD" and resid 56 and name HB1 ))
ASST (( 8122)
(( segid "BTD" and resid 63 and name HA ))
(( segid "BTD" and resid 62 and name HB1 ))
OR (( 8121)
3.400 1.400 1.400 peak 8122 weight
(( segid "BTD" and resid 14 and name HD21)
(( segid "BTD" and resid 18 and name HD1 ))
ASST (( 8492)
(( segid "BTD" and resid 18 and name HD21)
(( segid "BTD" and resid 14 and name HD1 ))
OR (( 8491)
3.400 1.400 1.400 peak 8492 weight
(( segid "BTD" and resid 18 and name HD21)
OR (( 8492)
(( segid "BTD" and resid 18 and name HD1 ))
(( segid "BTD" and resid 22 and name HD11)
ASST (( 8542)
(( segid "BTD" and resid 22 and name HD11)
(( segid "BTD" and resid 44 and name HD1 ))
OR (( 8541)
3.400 1.700 1.700 peak 8542 weight
(( segid "BTD" and resid 22 and name HD11)
OR (( 8542)
(( segid "BTD" and resid 22 and name HD11)
(( segid "BTD" and resid 22 and name HD1 ))
OR (( 8542)
(( segid "BTD" and resid 22 and name HD1 ))
(( segid "BTD" and resid 22 and name HD1 ))
ASST (( 8662)
(( segid "BTD" and resid 56 and name HD21)
(( segid "BTD" and resid 23 and name HD21)
OR (( 8661)
3.400 1.200 1.200 peak 8662 weight
(( segid "BTD" and resid 56 and name HD21)
(( segid "BTD" and resid 25 and name HD21)
ASST (( 8712)
(( segid "BTD" and resid 56 and name HD11)

```

11





13

14







81





20

21

[illegible]

23













Table 4

---

Hydrogen Bonding Restraints

---

## !Helix Z

assign (residue 19 and name HN )	(residue 15 and name O )	1.80	0.0	0.40
assign (residue 19 and name N )	(residue 15 and name O )	2.80	0.30	0.40
assign (residue 22 and name HN )	(residue 18 and name O )	1.80	0.0	0.40
assign (residue 22 and name N )	(residue 18 and name O )	2.80	0.30	0.40
assign (residue 23 and name HN )	(residue 19 and name O )	1.80	0.0	0.40
assign (residue 23 and name N )	(residue 19 and name O )	2.80	0.30	0.40
assign (residue 24 and name HN )	(residue 20 and name O )	1.80	0.0	0.40
assign (residue 24 and name N )	(residue 20 and name O )	2.80	0.30	0.40
assign (residue 25 and name HN )	(residue 21 and name O )	1.80	0.0	0.40
assign (residue 25 and name N )	(residue 21 and name O )	2.80	0.30	0.40

## !Helix B

assign (residue 75 and name HN )	(residue 71 and name O )	1.80	0.0	0.40
assign (residue 75 and name N )	(residue 71 and name O )	2.80	0.30	0.40
!assign (residue 77 and name HN )	(residue 73 and name O )	1.80	0.0	0.40
!assign (residue 77 and name N )	(residue 73 and name O )	2.80	0.30	0.40
assign (residue 78 and name HN )	(residue 74 and name O )	1.80	0.0	0.40
assign (residue 78 and name N )	(residue 74 and name O )	2.80	0.30	0.40
assign (residue 79 and name HN )	(residue 75 and name O )	1.80	0.0	0.40
assign (residue 79 and name N )	(residue 75 and name O )	2.80	0.30	0.40
!assign (residue 80 and name HN )	(residue 76 and name O )	1.80	0.0	0.40
!assign (residue 80 and name N )	(residue 76 and name O )	2.80	0.30	0.40
assign (residue 81 and name HN )	(residue 77 and name O )	1.80	0.0	0.40
assign (residue 81 and name N )	(residue 77 and name O )	2.80	0.30	0.40
assign (residue 82 and name HN )	(residue 78 and name O )	1.80	0.0	0.40
assign (residue 82 and name N )	(residue 78 and name O )	2.80	0.30	0.40

## !Helix C

assign (residue 102 and name HN )	(residue 98 and name O )	1.80	0.0	0.40
assign (residue 102 and name N )	(residue 98 and name O )	2.80	0.30	0.40
assign (residue 103 and name HN )	(residue 99 and name O )	1.80	0.0	0.40
assign (residue 103 and name N )	(residue 99 and name O )	2.80	0.30	0.40
assign (residue 104 and name HN )	(residue 100 and name O )	1.80	0.0	0.40
assign (residue 104 and name N )	(residue 100 and name O )	2.80	0.30	0.40
assign (residue 105 and name HN )	(residue 101 and name O )	1.80	0.0	0.40
assign (residue 105 and name N )	(residue 101 and name O )	2.80	0.30	0.40



[illegible]













ATOM	1947	H22	LVS	118	4.432	9.001	-5.038	1.00	0.00
ATOM	1948	C3	LVS	118	3.021	10.506	-5.527	1.00	0.00
ATOM	1949	H21	LVS	118	3.175	10.140	-6.531	1.00	0.00
ATOM	1950	H23	LVS	118	3.077	11.585	-5.522	1.00	0.00
ATOM	1951	CD	LVS	118	1.462	10.081	-5.053	1.00	0.00
ATOM	1952	H21	LVS	118	1.422	9.100	-5.449	1.00	0.00
ATOM	1953	H22	LVS	118	1.437	10.044	-5.374	1.00	0.00
ATOM	1954	C2	LVS	118	-0.025	11.347	-4.464	1.00	0.00
ATOM	1955	H21	LVS	118	1.048	11.322	-5.942	1.00	0.00
ATOM	1956	H22	LVS	118	-0.324	10.443	-6.543	1.00	0.00
ATOM	1957	ME	LVS	118	-0.724	9.558	-6.180	1.00	0.00
ATOM	1958	H21	LVS	118	0.214	10.234	-7.408	1.00	0.00
ATOM	1959	H22	LVS	118	-1.097	11.059	-5.778	1.00	0.00
ATOM	1960	C	LVS	118	-0.752	11.322	-5.942	1.00	0.00
ATOM	1961	OT1	LVS	118	4.769	13.177	-4.391	1.00	0.00
ATOM	1962	OT2	LVS	118	4.901	11.586	-2.445	1.00	0.00
END									





3











4.534	5.217	1.00	0.00	BFD ATOM	1945	CO1	11E	116	4.562
0.615	4.373	1.00	0.00	BFD ATOM	1946	HD11	11E	116	3.870
0.073	5.375	1.00	0.00	BFD ATOM	1947	HD12	11E	116	4.240
0.086	4.612	1.00	0.00	BFD ATOM	1948	HD13	11E	116	5.356
1.751	6.053	1.00	0.00	BFD ATOM	1949	C	11E	116	4.802
0.544	8.154	1.00	0.00	BFD ATOM	1951	N	AS9	117	4.323
2.465	5.014	1.00	0.00	BFD ATOM	1952	HM	AS9	117	4.079
1.454	8.874	1.00	0.00	BFD ATOM	1953	CA	AS9	117	5.805
1.482	10.275	1.00	0.00	BFD ATOM	1954	HA	AS9	117	5.197
0.609	10.018	1.00	0.00	BFD ATOM	1955	CH	AS9	117	4.967
0.267	10.220	1.00	0.00	BFD ATOM	1956	HM	AS9	117	4.384
0.596	4.976	1.00	0.00	BFD ATOM	1957	H22	AS9	117	3.723
0.555	10.876	1.00	0.00	BFD ATOM	1958	CO	AS9	117	2.897
1.488	10.787	1.00	0.00	BFD ATOM	1959	CO1	AS9	117	3.572
0.421	11.641	1.00	0.00	BFD ATOM	1960	CO2	AS9	117	7.006
0.882	12.208	1.00	0.00	BFD ATOM	1961	C	AS9	117	6.846
2.897	10.725	1.00	0.00	BFD ATOM	1962	O	AS9	117	8.206
2.384	9.879	1.00	0.00	BFD ATOM	1963	HM	LVS	118	9.416
1.613	11.488	1.00	0.00	BFD ATOM	1964	CA	LVS	118	10.248
2.063	10.968	1.00	0.00	BFD ATOM	1965	HA	LVS	118	9.315
2.220	12.889	1.00	0.00	BFD ATOM	1967	CH	LVS	118	8.436
2.199	11.949	1.00	0.00	BFD ATOM	1968	H21	LVS	118	10.252
3.656	12.895	1.00	0.00	BFD ATOM	1969	H22	LVS	118	7.818
3.719	12.296	1.00	0.00	BFD ATOM	1971	HO1	LVS	118	6.594
3.946	13.111	1.00	0.00	BFD ATOM	1972	HO2	LVS	118	9.857
4.808	12.327	1.00	0.00	BFD ATOM	1973	CO	LVS	118	10.477
4.872	13.133	1.00	0.00	BFD ATOM	1974	HO3	LVS	118	10.468
4.078	11.613	1.00	0.00	BFD ATOM	1976	CE	LVS	118	9.208
5.425	12.902	1.00	0.00	BFD ATOM	1977	H21	LVS	118	8.502
6.386	12.369	1.00	0.00	BFD ATOM	1978	H22	LVS	118	8.487
6.652	10.943	1.00	0.00	BFD ATOM	1980	H21	LVS	118	10.725
6.458	11.316	1.00	0.00	BFD ATOM	1981	H22	LVS	118	11.161
7.656	11.096	1.00	0.00	BFD ATOM	1982	H23	LVS	118	9.989
6.459	9.921	1.00	0.00	BFD ATOM	1983	C	LVS	118	10.206
0.441	12.023	1.00	0.00	BFD ATOM	1984	OT1	LVS	118	9.657
0.302	12.459	1.00	0.00	BFD END	1985	OT2	LVS	118	9.047
									10.436

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**